A Publication of the American Psychological Association

# Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

Articles

Own Attitude as an Anchor in Equal-Appearing Intervals ..... 85
*Harry S. Upshaw*

Theory and Experiment on the Measurement of Approach-Avoidance Conflict ..... 97
*Seymour Epstein and Walter D. Fenz*

Mediation and Associative Facilitation in Neurotic, Psychotic, and Normal Subjects ..... 113
*Peter J. Lang and Kenneth Lunde*

A Dissonance Theory Approach to Defensive Projection ..... 121
*Dana Bramel*

The Influence of Hunger on the Learning and Recall of Food Related Words ..... 129
*Seymour Epstein and Herbert Lewis*

Values Expressed in American Children's Readers: 1800–1950 ..... 136
*Richard deCharms and Gerald H. Moeller*

The Effects of Providing Knowledge of Results upon the Perceptual Defense Effect ..... 143
*Michael J. Goldstein and Samuel A. Himmelfarb*

Critiques and Notes

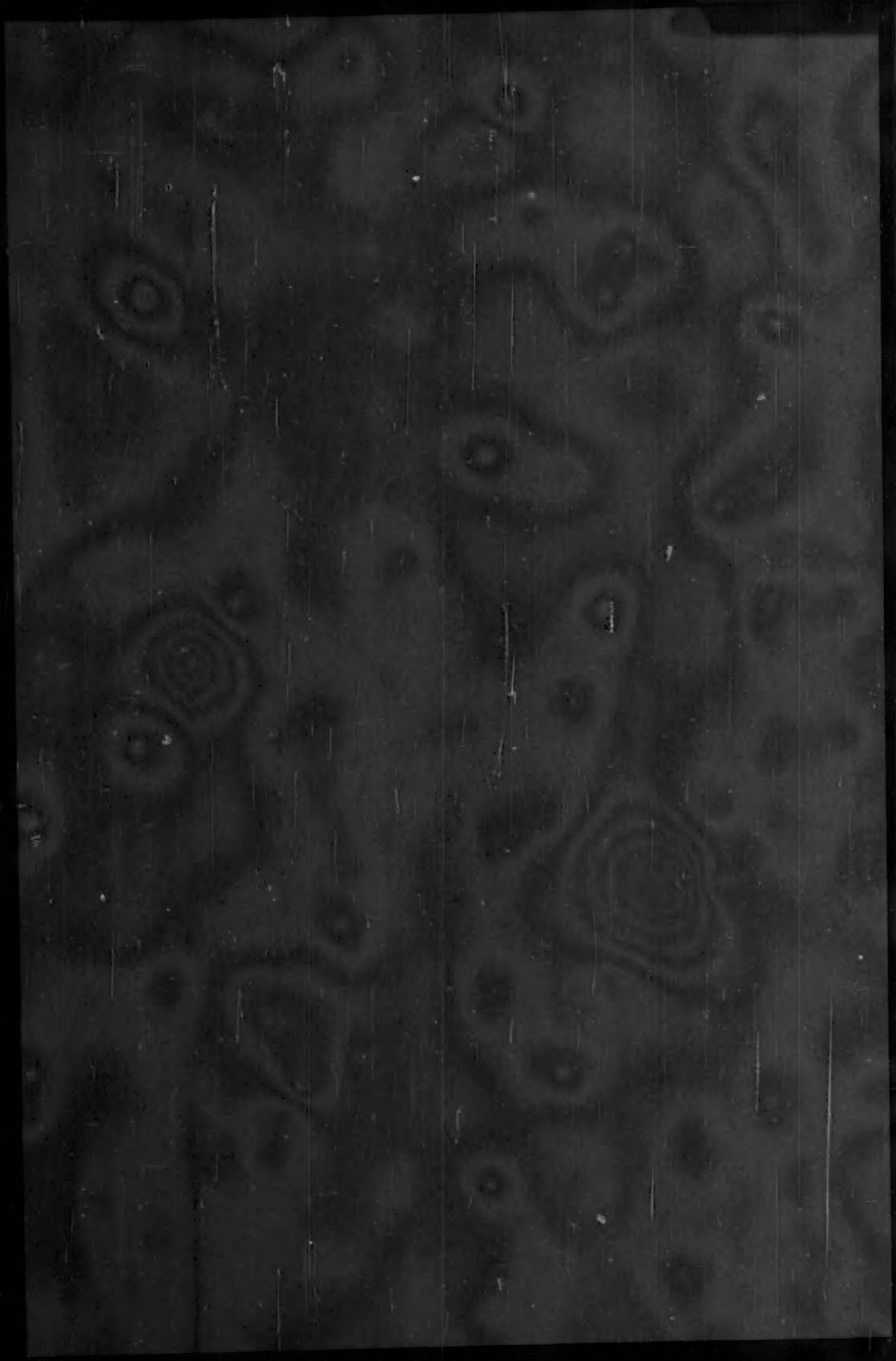
Seating Arrangement and Leadership Emergence ..... 148
*Lloyd T. Howells and Selwyn W. Becker*

The Effects of Patients' Relationships with Peers and Staff on Their Psychiatric Treatment Programs ..... 151
*E. Robert Sinnett and Dana B. Hanford*

Birth Order, Need Achievement, and Conformity ..... 155
*Edward E. Sampson*

The Recall of Dreams: Its Relation to Repression and Cognitive Control 160
*Frank M. Lachmann, Benjamin Laplin, and Norman S. Houlihan*

Volume 64 FEBRUARY 1962 Number 2



Volume 64 February 1962 Number 2

# Journal of Abnormal and Social Psychology

Editor: DANIEL KATZ, *University of Michigan*

Associate Editor: ROBERT B. ZAJONC, *University of Michigan*

CONSULTING EDITORS

JOSEPH B. ADELSON
*University of Michigan*

GORDON W. ALLPORT
*Harvard University*

NORMAN H. ANDERSON
*University of California, Los Angeles*

LEONARD BERKOWITZ
*University of Wisconsin*

JACK W. BREHM
*Duke University*

URIE BRONFENBRENNER
*Cornell University*

DONALD R. BROWN
*Bryn Mawr College*

DONALD T. CAMPBELL
*Northwestern University*

LAUNOR F. CARTER
*System Development Corporation*

DWIGHT W. CHAPMAN, JR.
*Vassar College*

MORTON DEUTSCH
*Bell Telephone Laboratories*

SEYMOUR FESHBACH
*University of Pennsylvania*

DONALD W. FISKE
*University of Chicago*

NATHANIEL L. GAGE
*University of Illinois*

NICHOLAS HOBBS
*Peabody College for Teachers*

MURRAY HORWITZ
*New York University*

IRVING L. JANIS
*Yale University*

HAROLD H. KELLEY
*University of California, Los Angeles*

HERBERT C. KELMAN
*Harvard University*

DANIEL J. LEVINSON
*Harvard Medical School*

DAVID C. MCCLELLAND
*Harvard University*

ROBERT B. MACLEOD
*Cornell University*

HAROLD L. RAUSH
*National Institute of Mental Health*

NEVITT SANFORD
*Stanford University*

IRVING SARNOFF
*Western Reserve University*

WILLIAM A. SCOTT
*University of Colorado*

JOSEPH VEROFF
*University of Michigan*

JUSTIN L. WEISS
*Massachusetts Mental Health Center*

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

The JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY is devoted to both abnormal and social psychology. Its emphasis is on basic research and theory rather than the techniques and arts of practice. Abnormal psychology is broadly defined to include papers contributing to fundamental knowledge of the pathology, dynamics, and development of personality or individual behavior, including deterioration with age and disease. Articles concerned with psychodiagnostic techniques are evaluated with respect to their contribution to an understanding of the psychological principles of diagnosis, those concerned with psychotherapy are judged in terms of their contribution to an understanding of the therapeutic process. Case reports are considered in terms of their heuristic value. Reports which serve to pose or to clarify important theoretical problems, or which promise to be useful in teaching, are often published. From the social area this JOURNAL gives preference to papers contributing to basic knowledge of interpersonal relations, and of group influences on the pathology, dynamics, and development of individual behavior.

The JOURNAL is issued monthly, two volumes per year. The subscription price per volume is $10.00, or $20.00 annually (foreign, $20.50). Single issues are $2.00.

Original manuscripts should be sent to the Editor:

DANIEL KATZ
University of Michigan
Box 1248
Ann Arbor, Michigan

---

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR
Acting Production Manager

Subscriptions, address changes, and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 Sixteenth Street, N.W.
Washington 6, D. C.

Changes of address must reach the Subscription Office by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Sts., Lancaster, Pa.
and 1333 Sixteenth Street N.W., Washington 6, D. C.

Second-class postage paid at Lancaster, Pa., and at additional mailing offices

Copyright © 1962 by the American Psychological Association, Inc.

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 85-96

# OWN ATTITUDE AS AN ANCHOR IN EQUAL-APPEARING INTERVALS [1]

HARRY S. UPSHAW

*University of North Carolina*

In recent years evidence has accumulated against Thurstone's (Thurstone & Chave, 1929) assumption that ratings of attitude statements in the method of equal-appearing intervals are independent of the attitudes of judges. Hovland and Sherif (1952) replicated Hinckley's (1932) study, arguing that the earlier finding of no effect on scale values due to judges' attitudes was traceable to Hinckley's use of a "carelessness criterion," which served to purge the judging group of people with extreme attitudes. Without the use of the disputed carelessness criterion, these investigators found that extreme judges tended to assign a large number of items to the categories at the end of the scale opposite to their own positions. Furthermore, a negative relationship was observed between the own positions of judges and the scale values assigned by them to midscale items. These results were interpreted as support for three hypotheses:

*Hypothesis I.* Judges with extremely pro or con attitudes will show a tendency to concentrate their placement of items into a small number of categories.

[1] This study was supported by a PHS grant (No. GN-5549) from the Division of Research Grants, United States Public Health Service.

The author is grateful to other members of the project staff of the "Study of the Essentials for Public Health Nursing Field Experience" for their aid in various phases of the research. In addition, appreciation is expressed to the following people for their particular help: John W. Thibaut, Peter Gumpert, and Bibb Latané, for stimulating discussions that led to the formulation of the problem and experimental design; W. J. Daniel and Paul Wood, for making arrangements for subjects; E. D. Hinckley, for making available the items used in his earlier study; Charles D. Ward, for numerous suggestions and other help in the analysis of data; Thomas Ostrom and Samuel Fillenbaum, for suggestions concerning the preparation of the manuscript. Finally, gratitude is expressed to the reviewer for this journal who made many suggestions for improvement of the report.

*Hypothesis II.* Judges with an extreme position and strong involvement will be highly discriminating in accepting items at their own end of the scale. They will correspondingly display a strong tendency to lump together statements at the end of the scale which they reject. The former tendency can be described as a raised *threshold of acceptance* and the latter as a lowered *threshold for rejection.*

*Hypothesis III.* A greater degree of displacement will occur for the "neutral" items and a smaller degree for the sharply defined pro and con statements at the extremes (Hovland & Sherif, 1952, p. 824).

The general flavor of the Hovland and Sherif report, and of Hypotheses II and III in particular, suggests an interpretation of the results in terms of a process of perceptual vigilance, whereby judges displace neutral items to the end of the scale opposite to their own position.

In a later paper, Hovland, Harvey, and Sherif (1957) presented a model to account for the effect of judges' attitudes on the evaluation of the position represented in a communication. Manis (1960) employed this model to explain the relationship he observed between the two variables. According to the model, which will be called here the "assimilation-contrast" model, items that a subject would personally reject are judged to be more distant from his own position than they really are (contrast), and items that represent positions only slightly discrepant from a subject's own attitude are judged to be closer to his own position than they really are (assimilation). It is interesting to note that the assimilation-contrast model makes allowance for both positive and negative relationships between judges' attitudes and item scale values. As a matter of fact, Manis observed a positive relationship for midscale items, whereas Hovland and Sherif found a negative relationship.

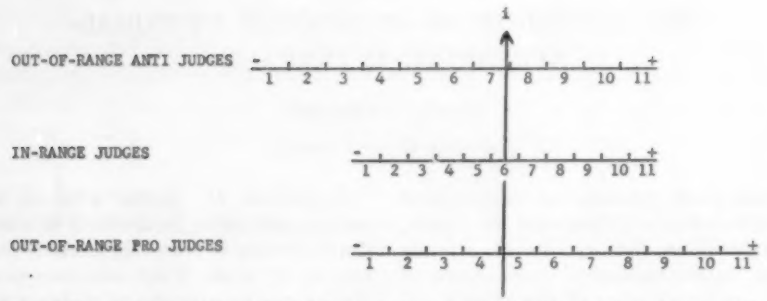



FIG. 1. Hypothetical scales adopted by out-of-range and in-range judges showing differences in judgments of an item *i* as a function of differences in these scales.

Still another theoretical orientation to the relationship between judges' attitudes and scale value is represented by the research of Segall (1959), who approached the problem as one of adaptation level. According to this model, a judge's attitude, in part, determines which stimulus he judges to be neutral, causing the point of neutrality, or adaptation level, to be displaced toward the judge's attitude. Because all absolute judgments are assumed to be based on the perceived distance of each item from the center of the reference scale, the judge's own position affects all of his judgments. Thus, a negative relationship between own position and scale value is hypothesized. Segall did not observe any systematic relationship between these variables, presumably because of an inadequate range of attitudes among his judges. Despite this failure to confirm the adaptation level hypothesis with respect to judges' attitudes, the model remains an important one to consider in the context of the equal-appearing intervals judging task.

A fourth theoretical model, which may be called "variable series," also applies to the effect of judges' attitudes on scale values. This model is based on the theoretical position of Volkmann (1951), who assumes that the stimulus range is the major determinant of the particular values assigned to items in a series. Judges whose own attitudes are included in the range of statements presented for judgment (i.e., *in-range* judges) are expected to divide the item range into segments corresponding to the categories specified in the experimental instructions. It may be assumed that the equal-appearing intervals judging task makes salient to a judge his own attitude, which functions for *out-of-range* judges as an anchor at one end of the scale. Thus, out-of-range judges subjectively divide a range corresponding to the experimental series augmented by their own attitudes. Accordingly, the scales adopted by in-range judges are narrower than those of out-of-range judges, leading to relatively finer discriminations among the items by in-range judges. Figure 1 illustrates the manner in which variable series accounts for a predicted negative relationship between own position and item judgments for an item that is in the middle of the series, and consequently at the middle of the scale adopted by in-range judges. The differences in item judgment illustrated in Figure 1 would be obliterated by extending the item range to include the positions of all judges, thus creating a constant series for all subjects.

The present experiment attempts to test differentially four theoretical interpretations of the relationship between judges' attitudes and equal-appearing intervals scale values. Inherent in the models of perceptual vigilance, assimilation-contrast, adaptation level, and variable series are two critical independent variables relevant to this relationship: the range of items presented for judgment, and the own positions of judges. The present design includes manipulation of both of these variables. Three experimental conditions were used: t (total series), in which it is assumed that the own positions of all judges fall within the item range, a + (series aborted at the pro end) and a − (series aborted at the anti end), in which it is assumed that the own

positions of many judges exceed the range. Within each of the experimental conditions, scale values of the items included in the series were determined for three groups of judges: A (anti judges), N (neutral judges), and P (pro judges). Nine sets of equal-appearing intervals item values were computed for the three groups of judges in the three experimental conditions. In the discussion to follow, the subgroups of judges and their sets of scale values are designated by a letter corresponding to the judges' attitudes, followed by a subscript representing an experimental condition (e.g., $A_t$ represents anti judges in the total series condition, $N_{a+}$ represents neutral judges in the condition in which the series was aborted at the pro end, etc.).

In terms of the present design, the four explanatory models are conceptualized as follows:

*Perceptual vigilance.* The critical independent variable in this model is personal involvement. Presumably, the more extreme a judge's own attitude, the greater his personal involvement, and the more likely he is to be threatened by the judging task. As a consequence of the threatening nature of the task, extremely pro and anti judges displace items in a direction opposite to their own attitudes. Neutral judges, having less personal involvement in the task, give more nearly accurate judgments than do extreme judges. Insofar as the absolute positions of the judges' attitudes is the factor accounting for any differences in scale values among the nine subgroups of the study, the manipulation of the item series has no relevance. With respect to the two independent variables of this study, perceptual vigilance predicts a main effect due to own attitude of judges, but no effect due to manipulation of the series, or to interaction. It is, however, assumed that the perceptual distortions of extreme judges are more likely to occur with ambiguous stimuli than with clear-cut, well-structured stimuli. Edwards (1946) has shown that midscale items on an equal-appearing intervals scale tend to be more ambiguous than extreme items. For this reason, it is expected that the predicted difference between judges, A > N > P (where ">" means "more favorable than") for all experimental conditions, occurs primarily in the judgments of midscale items.

*Assimilation-contrast.* In explaining the influence of judges' attitudes on scale values, this model assumes that each judge displaces, toward his own position, items that are only slightly removed from those he would endorse, and displaces, away from his position, items that he would reject. It is not feasible within the present design to compute latitudes of acceptance and rejection for each judge. However, on the assumption that pro judges would reject the most anti items and that anti judges would reject the most pro items, the assimilation-contrast model implies a significant main effect due to the variable of judges' attitudes. Furthermore, if assimilation (displacement toward own position) is observed at all, it must be for items that are closer to own position than any items displaying the contrast phenomenon (displacement away from own position). Like the perceptual vigilance model, assimilation-contrast predicts no effect due to manipulation of the item series, and no effect due to interaction between judges' attitudes and item series.

*Adaptation level.* This model assumes that item judgments depend on the perceived distance of each item above or below the adaptation level (point of neutrality). A judge's adaptation level is the result of three types of influence: a weighted average of the stimulus values of items in the series; an effect due to anchor stimuli or other contextual factors; and a residual effect due to pre-experimental factors (Michels & Helson, 1949). The attitudes of judges may be conceived, as in Segall (1959), to contribute to a residual effect, or to serve an anchor function. With either conception, it is assumed that the adaptation level of each judge is displaced toward his own attitude, producing a displacement of item judgments in the opposite direction. A main effect due to judges' attitudes is predicted, in the order: A > N > P. Because it produces variation in the weighted average of the stimulus values, the manipulation of the item series leads to predicted shifts in adaptation level toward the anti end of the scale in the a + condition, and toward the pro end in the a − condition. As a result of these shifts in adaptation level, a difference due to

experimental conditions is predicted, with item values in the order: $a+ > t > a-$. Hence, unlike perceptual vigilance and assimilation-contrast, which predict a single main effect, adaptation level predicts main effects due to item series, as well as to judges' attitudes.

*Variable series.* The critical factor in this interpretation is whether the item series includes a judge's own position. In-range judges are assumed to adopt reference scales defined by the items at either extreme of the item series, whereas out-of-range judges adopt scales anchored by their own attitudes at one end, and the most extreme item in the series at the other. It is intended that the series in the t condition covers the range of judges' attitudes. Therefore, the prediction is made that $A_t = N_t = P_t$. With the most favorable items removed in the a + condition, the scales adopted by P judges are predictably wider than those of N and A judges due to differential anchoring at the pro end. The result of the varying widths of reference scales is variation in item judgments as a function of own position, a given item receiving a lower value on a scale that extends farther at the pro end. Therefore, the predicted order of item judgments is $A_{a+} \geqslant N_{a+} > P_{a+}$. ($A_{a+}$ is predicted to be equal to or greater than $N_{a+}$, depending on whether the experimentally aborted series covers the positions of both groups.) By similar reasoning for the a − condition, the predicted order of item judgments is $A_{a-} > N_{a-} \geqslant P_{a-}$, due to the extension of the scales adopted by A judges in the anti direction. Phrased another way, it is the manipulation of the series that produces a difference due to own positions. Thus, variable series predicts a main effect due to judges' attitudes, as do the perceptual vigilance, assimilation-contrast, and adaptation level models. It also predicts a main effect due to manipulation of the series, as does adaptation level. The distinctive prediction for variable series, however, is an interaction between own position and item series. In particular, the prediction is that, because A, N, and P judges shift differentially in scale width as a function of the experimental conditions, the difference between the t and a + conditions is smaller for P judges than for A and N judges, and the difference between the t and a − conditions is smaller for A judges than for P and N judges.

## METHOD

### *Procedure*

In order to provide maximum comparability with previous research, the present study is essentially a replication of the Hinckley (1932) and Hovland and Sherif (1952) studies. The major difference between the present design and that of the earlier studies is that the range of the item series, as well as judges' attitudes, is treated as an independent variable. The manipulation of item range was accomplished, as in Fehrer (1952), by preparing three sets of attitude items. One set is identical to that employed by Hinckley and by Hovland and Sherif, consisting of all of the 114 statements assembled by Hinckley relating to the social position of the Negro. The second set of items represents an aborted range of statements with the 78 most favorable items removed on the basis of the original Hinckley scale values. (The decision concerning the number of items to be removed was an arbitrary one.) The third set has the 78 most unfavorable items removed.

Five hundred students of introductory psychology at the University of North Carolina served as judges, partitioned at random into three groups, each group receiving one of the sets of items. All judges received the following instructions, adapted from Hinckley (1932):

1. You have been solicited as a subject in the construction of a scale for measuring attitude toward the Negro.

2. This envelope contains a series of opinions about the *social position of the Negro*. Your task is to sort these opinions into 11 piles.

3. To aid you in this task, you will find enclosed eleven slips with Roman numbers on them. Please lay these out in the following order: I, II, III, IV, V, VI, VII, VIII, IX, X, XI. *On slip I*, put those statements which place the Negro in the lowest social position. *On slip XI*, put those statements which place the Negro in the highest social position. In like manner, place the other statements on slips II, III, IV, etc., as they give the Negro a higher and higher social position. You will then have all of the opinions arranged in eleven piles, corresponding to eleven steps in the social ladder.

4. Do not try to get the same number in each pile. They are not evenly distributed.

5. You can probably finish the task in less than an hour.

6. Kindly place your name and classification on enclosed slip.

7. When you have finished, please clip all statements which you have placed in each pile together, with the proper Roman number slip on top. Replace the eleven piles thus clipped into the envelope.

After the sorting task, all judges completed the Murphy-Likert Scale of Attitude toward the Negro (Likert, 1932). On the basis of the distribution of their scores, three groups were constituted, with those in the most unfavorable quintile forming the A group, those in the middle quintile forming the N group, and those in the most favorable quintile forming the P group. The Murphy-Likert scale has a maximum score range of 20–70. The actual range of scores in this study was 38–70. The anti-Negro quintile was defined in terms of scores lower than 55; the neutral quintile, scores between 59 and 62; and the pro-Negro quintile, scores higher than 65. Only A, N, and P judges were of concern in this study. After eliminating judges in the second and fourth Likert quintiles, 321 subjects remained.

### *Data Analysis*

Nine sets of equal-appearing intervals values were computed for the various subgroups corresponding to judges' attitude and the range of items judged. The subgroups varied in size from 29 in the $A_t$ condition to 45 in the $P_t$ condition. "Control scale values" were computed, based on the judgments of the entire group of 175 subjects who participated in the t condition, including those in the second and fourth Likert quintiles, as well as those in conditions $A_t$, $N_t$, and $P_t$. On the basis of the control scale values, the nine sets of values for the experimental subgroups were partitioned into subsets of extremely pro items (13 items with scale values between 9.50 and 11.49), moderately pro (18 items with values between 7.50 and 9.49), midscale (30 items between 4.50 and 7.49), moderately anti (24 items between 2.50 and 4.49), and extremely anti items (29 items between 0.50 and 2.49).[2]

The hypotheses for this study, derived from the four theoretical models, are phrased in terms of main effects and interaction. This formulation suggests a two-way analysis of variance as an appropriate statistical procedure. However, it should be noted that one of the four models, variable series, is based upon consideration of the effect of the two independent variables on the width of the reference scales adopted by judges. Whereas variations in scale width lead to predictable differences in the modal judgments of items, such variations also lead to predictable effects on the dispersion of judgments, which may render untenable the homogeneity of variance assumptions of the analysis of variance statistical models. As a matter of fact, preliminary analyses of some of the present data by means of the method of successive intervals, which permits solution of the item discriminal dispersions, indicate that there are significant effects on item variance due to judges' attitudes, item series, and interaction between these variables.[3] To overcome this difficulty, a nonparametric analogue of the analysis of variance was used, employing a statistic called $\chi^2_r$, the distribution of which approaches that of $\chi^2$ as the number of replications (in this case, items) becomes indefinitely large (Friedman, 1937, 1940; Wilcoxon, 1949).

In order to test the main effect due to judges' attitudes, it was necessary to determine for A, N, and P judges scale values that were not systematically influenced by the manipulation of item series. To this end, the medians of the values assigned to each item under the three experimental conditions were computed for the three groups of judges. On the basis of these medians the three judging groups were ranked, item-by-item, for extremely pro, moderately pro, midscale, moderately anti, and extremely anti items. For each of the item subsets, a sum of ranks was computed, and these sums for A, N, and P judges were tested for equality by means of $\chi^2_r$. An over-all test of the significance of own attitude was made by summing the values of $\chi^2_r$ and the degrees of freedom for the five item subsets. The evaluation of the effect of the manipulation of item series was conducted in a similar fashion. For each item, the median scale values assigned in the t condition by A, N, and P judges were determined, as were the medians for the a+ and a− conditions. The experimental conditions were ranked, item-by-item, for each of the subsets of items. Again, $\chi^2_r$ was used to test the hypothesis that the sum of ranks assigned to the experimental conditions are equal. Summing the $\chi^2_r$'s over the item subsets provided an over-all test of the effect due to item series.

There are two components of the interaction effect predicted from the variable series model. First, the differences between judgments in the a+ and t conditions were expected to be greater for A and N judges than for P judges. This expectation is based upon consideration of the relatively narrower scale of A and N judges in the a+ condition, with an accompanying displacement of judgments in the direction of the aborted end of the scale. To test this hypothesis, the scale values assigned by A, N, and P judges in the t condition were subtracted, item-by-item, from those obtained in the a+ condition. The A, N, and P judges were ranked according to these differences, the rank sums were evaluated by $\chi^2_r$ for each item subset, and an over-all test was made on the basis of the sum of the $\chi^2_r$'s for the five item subsets. Similarly, the differences between judgments in the a− and t conditions were expected to be greater for N and P judges than for A judges, due to the maintenance of a relatively wide scale by the latter in this aborted series condition. As a result of their narrower scales in the a− condition, N and P judges were

[2] An 11-page document listing the items constituting the three experimental series, the obtained scale and $Q$ values for the nine subgroups, and the control scale and $Q$ values is available through the American Documentation Institute. Order Document No. 6942 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress; Washington 25, D. C., remitting in advance $1.75 for microfilm or $2.50 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

[3] The data will be the subject of a later report.

TABLE 1
MEAN ITEM VALUES

| Experimental condition | Attitude of judges | Item subset | | | | |
|---|---|---|---|---|---|---|
| | | Extremely pro | Moderately pro | Midscale | Moderately anti | Extremely anti |
| *t* | A | 9.87 | 8.22 | 6.15 | 3.98 | 2.09 |
| | N | 10.20 | 8.56 | 6.23 | 3.26 | 1.62 |
| | P | 10.33 | 9.03 | 5.87 | 3.19 | 1.53 |
| a + | A | — | — | 6.64 | 4.33 | 2.46 |
| | N | — | — | 6.63 | 3.73 | 1.79 |
| | P | — | — | 5.83 | 2.93 | 1.40 |
| a − | A | 9.79 | 8.27 | 6.61 | 3.75 | — |
| | N | 9.72 | 8.30 | 5.89 | 2.93 | — |
| | P | 9.98 | 8.61 | 5.40 | 2.53 | — |
| Control | | 10.14 | 8.63 | 6.04 | 3.37 | 1.66 |

Note.—A dash indicates that an experimental manipulation rendered impossible the collection of data otherwise appropriate for a cell in the table. Dashes appear for both extremely pro and moderately pro items in the a + condition and for only the extremely anti items in the a − condition because of the generally lower values obtained in the present study as compared with those obtained by Hinckley (1932). Only four of the anti items removed in the a − condition, on the basis of the Hinckley data, received moderately anti values in the present study.

expected to displace their judgments toward the aborted, anti end of the scale. Accordingly, scale values obtained for each item in the t condition were subtracted from the corresponding values in the a − condition, and the differences evaluated by the ranking procedure for A, N, and P judges. The $\chi^2_r$'s computed for each item subset were summed to obtain an over-all evaluation of differences between a − and t conditions. Finally, the $\chi^2_r$'s corresponding to the two components of the interaction, a − minus t and a + minus t, were summed to provide a means of testing the total effect due to interaction.

## RESULTS

Table 1 presents the mean scale values assigned to items of each subset by A, N, and P judges under the three experimental conditions. Table 2 summarizes the results of the nonparametric tests of significance of the effects due to own attitude and item series, and Table 3 summarizes the results of the tests of interaction. The relative sizes of the rank sums listed in Tables 2 and 3 reflect the order of scale values assigned by the judges participating in the designated comparisons, with lower sums generally corresponding to lower scale values. From Tables 2 and 3 it can be seen that, considering all items, the effects due to own attitude, item series, and interaction are all highly significant. Of the four explanatory models, only variable series accounts for all three of these effects. However, not all of the analyses for separate item subsets are consistent with this

TABLE 2
SUMS OF RANKS AND TESTS OF SIGNIFICANCE FOR MAIN EFFECTS

| Item subset | Attitude of judges | | | $\chi^2_r$ | *df* | Experimental condition | | | $\chi^2_r$ | *df* |
|---|---|---|---|---|---|---|---|---|---|---|
| | A | N | P | | | t | a + | a − | | |
| Extremely pro | 19.0 | 23.0 | 36.0 | 12.15*** | 2 | 23.5 | — | 15.5 | 4.92* | 1 |
| Moderately pro | 30.0 | 31.0 | 47.0 | 10.11** | 2 | 31.5 | — | 22.5 | 4.50* | 1 |
| Midscale | 65.0 | 67.0 | 42.0 | 13.31*** | 2 | 62.5 | 69.5 | 42.0 | 14.09*** | 2 |
| Moderately anti | 71.0 | 47.0 | 26.0 | 42.25*** | 2 | 48.0 | 66.0 | 30.0 | 27.00*** | 2 |
| Extremely anti | 81.0 | 52.0 | 29.0 | 50.30*** | 2 | 31.5 | 49.5 | — | 12.00*** | 1 |
| All items | | | | 128.12*** | 10 | | | | 62.51*** | 7 |

Note.—A dash indicates that an experimental manipulation rendered impossible thec ollection of data otherwise appropriate for a cell in the table.
* $p < .050$, two-tailed test.
** $p < .010$, two-tailed test.
*** $p < .005$, two-tailed test.

TABLE 3

SUMS OF RANKS AND TESTS OF SIGNIFICANCE FOR INTERACTION

| Component of interaction | Item subset | Attitude of judges | | | $\chi^2_r$ | $df$ |
|---|---|---|---|---|---|---|
| | | A | N | P | | |
| a − minus t | Extremely pro | 32.0 | 23.0 | 23.0 | 4.15* | 2 |
| | Moderately pro | 43.5 | 34.5 | 30.0 | 5.25** | 2 |
| | Midscale | 77.0 | 51.0 | 46.0 | 19.10**** | 2 |
| | Moderately anti | 51.5 | 52.0 | 40.5 | 3.52* | 2 |
| | Extremely anti | — | — | — | — | — |
| | All items | | | | 32.02**** | 8 |
| a + minus t | Extremely pro | — | — | — | — | — |
| | Moderately pro | — | — | — | — | — |
| | Midscale | 69.0 | 57.0 | 48.0 | 7.66*** | 2 |
| | Moderately anti | 51.0 | 62.0 | 31.0 | 20.58**** | 2 |
| | Extremely anti | 64.0 | 61.0 | 37.0 | 16.22**** | 2 |
| | All items | | | | 44.46**** | 6 |
| Combined | Extremely pro | | | | 4.15* | 2 |
| | Moderately pro | | | | 5.25** | 2 |
| | Midscale | | | | 26.76**** | 4 |
| | Moderately anti | | | | 24.10**** | 4 |
| | Extremely anti | | | | 16.22**** | 2 |
| | All items | | | | 76.48**** | 14 |

Note.—A dash indicates that an experimental manipulation rendered impossible the collection of data otherwise appropriate for a cell in the table.

* $p < .100$, one-tailed test.
** $p < .050$, one-tailed test.
*** $p < .025$, one-tailed test.
**** $p < .001$, one-tailed test.

conclusion. For example, although a significant difference was found among A, N, and P judges for moderately pro and extremely pro items, the direction of the difference is opposite to that predicted not only by variable series, but also by perceptual vigilance and adaptation level. The finding that higher (more pro) values were assigned to these items by pro judges is consistent with the assimilation-contrast model. However, for two reasons, caution must be exercised in claiming support for the model on the basis of this finding. First, a contrary result, that pro judges assigned lower values to these items, would have been equally consistent with the model (interpretable as contrast, rather than assimilation). Second, the significant effects due to item series and interaction observed for these items were not anticipated on the basis of the model. Furthermore, as will be discussed in the next section, there are supplementary data suggesting that some of the items showing the apparent assimilation effect are well within the latitudes of acceptance of P judges, whereas the Hovland et al. (1957) model specifies assimilation of items somewhat removed from this region of the scale.

The analyses of the effect of own position on judgments of midscale, moderately anti, and extremely anti items are consistent with the predictions from all four models (interpretable as contrast effects in terms of the assimilation-contrast model).

With respect to the experimental manipulation of item range, the results reported in Table 2 are consistent with the predictions from both adaptation level and variable series for each of the item subsets. The demonstration of a significant effect due to the item series is not, in itself, relevant to scale construction by the method of equal-appearing intervals, which conventionally involves a standard series for all judges. If, as conceptualized by the variable series model, the attitudes of out-of-range judges contribute to the definition of their reference scales, the item series assumes importance by virtue of

the interaction between it and own attitudes.

Variable series is the only model of the four considered in this research which specifies an interaction between the two independent variables. It, furthermore, specifies the direction of the interaction. Therefore, the data reported in Table 3 constitute a critical test of that model. By reference to the table, it can be seen that, over-all, the interaction effect is highly significant. The obtained $\chi^2_r$'s for the a − minus t component for extremely pro and moderately anti items fall short of the values conventionally adjudged statistically significant. In both of these cases, however, the direction of the differences in rank sum for A, N, and P judges is that predicted by variable series. (It should be noted that, while the $\chi^2_r$ for extremely pro items in the combined interaction also falls short of the acceptable level, this is due entirely to the a − minus t component, since the combined term for extremely pro items was comprised only of that component.) Whether or not allowances are made for the two discrepant cases, collectively the analyses reported in Table 3 provide evidence in support of the variable series prediction that judges' attitudes and the item series interact in their influence on equal-appearing intervals scale values. This conclusion can be drawn even though it is apparent in the analyses of the main effect due to judges' attitudes that the variable series model cannot account for the data obtained with extremely pro and moderately pro items.

## SUPPLEMENTARY DATA

The variable series model was formulated on the basis of evidence in the psychophysical literature that a stimulus not present in the series can serve an end anchor function, defining the upper or lower end of the reference scale (e.g., Hunt & Volkmann, 1937). As envisaged in the present design, the total series would contain items more extreme at either end than any judge's own attitude. With the removal of the most pro items for the a + condition, it was assumed that the upper end anchors for pro judges would be their own attitudes, whereas the upper end anchors for neutral and anti judges would be either their own attitudes or the most favorable item in the aborted series. In any event, it seemed reasonable to assume that the reference scales adopted by A, N, and P judges would, because of differential end anchoring, differ in width more in the a + condition than in the t condition. Furthermore, it was assumed that there would be greater similarity between the scales of $P_t$ and $P_{a+}$ judges than between those of either $A_t$ and $A_{a+}$ judges, or $N_t$ and $N_{a+}$ judges. By the same reasoning, greater similarity was assumed between the scales of $A_t$ and $A_{a-}$ judges than between either $P_t$ and $P_{a-}$ judges, or $N_t$ and $N_{a-}$ judges.

After the decision to use the Hinckley (1932) items in this research, it was deemed inadvisable to assess own position by means of the same items. Memory of the rating previously assigned to an item could, conceivably, influence the likelihood of a subject's endorsing the item. Conversely, memory of an endorsement or rejection of an item could alter the rating of the item. Perhaps it would have been wise to have removed from the item pool a subset of items with which own position could have been measured. It was decided, however, to assess the attitudes of judges by means of the Murphy-Likert scale. Likert scales are ordinal scales, permitting only an ordering of subjects according to their attitudes. This type of measure is adequate in the present design for the two aborted range conditions. However, minimal validity of the experimental manipulations for testing variable series depends upon two assumptions. First, stimulus anchoring (as opposed to attitude anchoring) is assumed in the t condition at the pro end for anti judges, and at the anti end for pro judges. Second, attitude anchoring is assumed at the pro end for pro judges in the a + condition, and at the anti end for anti judges in the a − condition.

To assess the validity of the experimental manipulations for providing a test of the variable series model, new data were collected. A questionnaire was administered to 201 students drawn from the same undergraduate psychology course that a year earlier supplied the judges in the main experiment. The questionnaire included a 20-item equal-appearing intervals scale based on the control values obtained in the main experiment. The

items were chosen to obtain maximum coverage of the attitude continuum with minimum $Q$ values. The scale appeared as the third part of the questionnaire, the first two parts being concerned with the judgment of items, none of which were included in the 20-item attitude scale. The following instructions were given, adapted from those used by Hovland et al. (1957) to determine latitudes of acceptance and rejection:

1. This part contains 20 statements concerning the social position of the Negro. This time, however, you are requested to indicate your own acceptance or rejection of the items as expressions of your own opinion.

2. Read all of the items carefully before you make any judgments.

3. When you have read all of the statements look at the table on the next page. This table has five headings: Most Acceptable, Also Acceptable, Neither Acceptable nor Objectionable, Also Objectionable, and Most Objectionable.

4. Your task is to:

a. Find the *one* item in this set of 20 items that is *most acceptable* from your point of view, and list the item number for that one statement under the Most Acceptable column in the table.

b. There may be other statements which you find *also acceptable* from your point of view. Put the item numbers for these statements (if there are any) under the Also Acceptable column of the table.

c. Now find the *one* statement which is *most objectionable* from your point of view, and enter that statement's number under the Most Objectionable column in the table.

d. There may be other statements which you find *also objectionable* from your point of view. List the numbers of these statements (if there are any) under the Also Objectionable column of the table.

e. In the last column, Neither Acceptable nor Objectionable, the numbers of the remaining items (if there are any) should be listed. These remaining statements should be statements that from your point of view are *neither acceptable nor objectionable.*

f. As a check to make sure that all 20 items have been listed in the table, count the item numbers listed there. If not all 20 have been listed, find the missing items and list their numbers in the appropriate columns.

5. As in previous parts, the items have identification numbers, but these, again, should be ignored.

The fourth part of the questionnaire was the Murphy-Likert scale, on the basis of which A, N, and P judges were selected by the same criteria used in the main experiment. Of the 201 judges, 47 were thus assigned to the A group, 45 to the N group, and 37 to the P group. For each judge in the three groups the range of acceptable items on the equal-appearing intervals scale was determined (whether recorded as "most acceptable" or "also acceptable"). Table 4 shows the lower and upper boundaries of acceptable items for the three groups. These boundaries are represented in the table by the cumulative frequency and cumulative proportion of judges in each group for whom any item of the scale is the lowest and highest valued item endorsed. It is interesting to note that, considering the most extremely anti items endorsed, for only 16 (a proportion of .34) of the A judges were these items in the region that was excluded in the a− condition. Approximately 25% of the items in the Hinckley (1932) series are in this region. Furthermore, more than one-fourth of the A judges endorsed items with values above 9.69. This information, together with the finding that all 37 of the P judges endorsed the most extremely pro item of the scale, suggests that the total Hinckley series does not adequately cover the range of attitudes of white judges selected from this southern university. If this is so, it must certainly be true that the Negro judges in the Hovland and Sherif (1952) study were all out-of-range judges.

Assuming that the student population that supplied data for Table 4 was the same as that participating in the main experiment, the table provides information on the basis of which the validity of the earlier experimental manipulations can be evaluated. If those manipulations were minimally adequate to provide a test of the variable series formulation, stimulus anchoring must have prevailed in the t condition for A judges at the pro end, and for P judges at the anti end. It is likely that the large number of A judges who endorsed the most pro item of the scale reflects an end effect, and that many of these people would have endorsed items that are even higher in scale value, if they had been offered. Although it appears likely that stimulus anchoring did not prevail in this situation for all judges, it is probable that it did prevail for perhaps as many as three-fourths of the A group. At the anti end, it seems reasonable to conclude that stimulus anchoring prevailed for all P judges.

TABLE 4

CUMULATIVE FREQUENCY AND CUMULATIVE PROPORTION OF MOST EXTREMELY ANTI AND MOST EXTREMELY PRO ITEMS ENDORSED BY A, N, AND P JUDGES

| Item scale value | Most anti item endorsed | | | | | | Most pro item endorsed | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A judges | | N judges | | P judges | | A judges | | N judges | | P judges | |
| | Cumulative frequency | Cumulative proportion | Cumulative frequency | Cumulative proportion | Cumulative frequency | Cumulative proportion | Cumulative frequency | Cumulative proportion | Cumulative frequency | Cumulative proportion | Cumulative frequency | Cumulative proportion |
| 10.52 | 47 | 1.00 | 45 | 1.00 | 37 | 1.00 | 47 | 1.00 | 45 | 1.00 | 37 | 1.00 |
| 10.18 | 47 | 1.00 | 45 | 1.00 | 37 | 1.00 | 36 | .77 | 6 | .13 | 0 | .00 |
| 9.69 | 47 | 1.00 | 45 | 1.00 | 37 | 1.00 | 35 | .74 | 5 | .11 | 0 | .00 |
| 9.23 | 47 | 1.00 | 45 | 1.00 | 37 | 1.00 | 10 | .21 | 0 | .00 | 0 | .00 |
| 8.53 | 47 | 1.00 | 45 | 1.00 | 37 | 1.00 | 9 | .19 | 0 | .00 | 0 | .00 |
| 8.33 | 47 | 1.00 | 44 | .98 | 37 | 1.00 | 8 | .17 | 0 | .00 | 0 | .00 |
| 7.57 | 47 | 1.00 | 43 | .96 | 33 | .89 | 6 | .13 | 0 | .00 | 0 | .00 |
| 7.27 | 47 | 1.00 | 41 | .91 | 23 | .62 | 6 | .13 | 0 | .00 | 0 | .00 |
| 6.86 | 46 | .98 | 37 | .82 | 18 | .49 | 6 | .13 | 0 | .00 | 0 | .00 |
| 5.98 | 46 | .98 | 37 | .82 | 17 | .46 | 4 | .09 | 0 | .00 | 0 | .00 |
| 5.71 | 46 | .98 | 35 | .78 | 13 | .35 | 3 | .06 | 0 | .00 | 0 | .00 |
| 5.18 | 45 | .96 | 32 | .71 | 12 | .32 | 1 | .02 | 0 | .00 | 0 | .00 |
| 4.56 | 44 | .94 | 32 | .71 | 12 | .32 | 0 | .00 | 0 | .00 | 0 | .00 |
| 4.34 | 39 | .83 | 21 | .47 | 6 | .16 | 0 | .00 | 0 | .00 | 0 | .00 |
| 3.65 | 32 | .68 | 14 | .31 | 3 | .08 | 0 | .00 | 0 | .00 | 0 | .00 |
| 3.15 | 30 | .64 | 14 | .31 | 3 | .08 | 0 | .00 | 0 | .00 | 0 | .00 |
| 2.55 | 18 | .38 | 7 | .16 | 3 | .08 | 0 | .00 | 0 | .00 | 0 | .00 |
| 2.16 | 16 | .34 | 6 | .13 | 3 | .08 | 0 | .00 | 0 | .00 | 0 | .00 |
| 1.63 | 9 | .19 | 0 | .00 | 1 | .03 | 0 | .00 | 0 | .00 | 0 | .00 |
| 1.14 | 4 | .09 | 0 | .00 | 0 | .00 | 0 | .00 | 0 | .00 | 0 | .00 |

Note.—Items in the first segment of the table were excluded in the a + condition. Those in the third segment were excluded in the a − condition.

A second assumption on which the validity of the manipulations depends is that attitude anchoring prevailed at the pro end for $P_{a+}$ judges, and at the anti end for $A_{a-}$ judges. Table 4 indicates that all P judges endorsed statements above the pro stimulus anchor of the a + series and that .34 of the A judges endorsed statements below the anti stimulus anchor of the a − series. Again it can be concluded that the criterion was probably met by the P judges. The criterion with respect to A judges was probably met by about one-third of the judges.

On the basis of the data relating to the two critical assumptions, it seems very likely that the experimental procedure was adequate for P judges. There are grounds for disputing the validity of the procedure for A judges, although it is probable that enough of these judges satisfied the criteria to account for the results obtained with A judges in the main experiment.

In the presentation of results of the main experiment it was noted that P judges assigned higher values to extremely pro and moderately pro items than did A judges. In terms of the assimilation-contrast model, this finding could represent assimilation on the part of P judges. Assimilation is assumed to occur at, or slightly beyond, the boundary of a judge's latitude of acceptance. The procedures used by Hovland et al. (1957) to determine latitudes of acceptance were employed in the collection of the data represented in Table 4. From the table, it can be seen that all P judges endorsed the item with scale value 10.52, and that they all endorsed other items with values of 8.33, or lower. To the extent that the subjects supplying these supplementary data are like those participating in the main experiment, assimilation by P judges should occur only for items with values equal to, or lower than, 8.33. The extremely pro items, however, are items between 9.50 and 11.50, whereas moderately

pro items have values between 7.50 and 9.49. Hence, the extremely pro items are well within the latitude of acceptance of P judges. This fact, together with other evidence already discussed, casts suspicion on the adequacy of assimilation-contrast, as presently formulated, to account for the data with respect to pro items; data which none of the other models can account for, either.

## DISCUSSION

The data of this study indicate that there is a differential effect of own position on equal-appearing intervals scale values according to the range of items presented for judgment. Of the four explanatory models considered, only variable series specifies such an interaction. However, the positive relationship between own position and scale value obtained in this study for pro items cannot be simply explained by variable series or any of the other models considered here. Perhaps own position is not the critical variable in this context, or perhaps it is not the only critical variable. One interpretation of the present experiment is that, when confronted with a judgmental task involving an item series which, for him, is aborted, a judge supplies anchors based on his expectation concerning what the series should be. The judgmental task may make salient to the judge various anchors at several points along the continuum, only one of which is his own attitude. If the series does not include items around these points, then, the self-supplied anchors may assume importance in determining judgmental behavior.

It may be objected that the choice of subjects did not permit an adequate test of the perceptual vigilance interpretation. Hovland and Sherif (1952) selected judges with extreme attitudes and known personal involvement in the race issue. The present study, on the other hand, used only judges of the type described by Hovland and Sherif as "average white judges." The "average white judges" in this study became extreme judges in some experimental conditions by virtue of the manipulation of the item series. Whether judges who are extreme in their "natural state" (presumably as defined by national norms) behave differently from those called extreme in this study is a problem now being investigated by Thomas Ostrom of this university. Pending Ostrom's results, the influence of truly extreme attitudes on judgments of the type considered here is a matter of conjecture. However, any hypothesis concerning the judgmental processes of involved, extreme judges should take into account the interaction effect obtained in this study between own attitude and the item series. Perceptual vigilance as presently formulated does not.

The assimilation-contrast model, similarly, does not consider the influence of the item series. In the experiment accompanying the statement of the model (Hovland et al., 1957), the evaluations of single, 15-minute oral communications were analyzed in terms of the model. Manis (1960) analyzed the judgments of 12 attitude statements. It could be that the model is appropriate for some conditions of item series, but not for others. Research directed to these limits of applicability is suggested by the findings of this study.

Although the data concerning the two main effects due to own attitude and item series are consistent with the adaptation level model, those concerning the interaction between these variables are not. On the surface, it is difficult to imagine how any influence on the point of perceived neutrality can be made to account for this interaction.

Of the four explanatory models, variable series is best supported by the data. Applied to the judgmental task of equal-appearing intervals, the variable series interpretation assumes, in effect, that the own attitudes of judges are an extraneous variable which acquires importance only when the item series is such that the positions of some judges are outside it. A reference scale, defined by end anchors corresponding to the most pro and anti items, is inherent in the series. If a judge has an own position that is outside the scale implied by the series, he does not adopt that scale. Instead, he adopts one that is defined by his own position serving as an end anchor at the aborted end of the continuum. Consequently, his scale is broader than that implied by the series, and his judgments accordingly reflect grosser discriminations than

those of his fellow judges who accept the series-defined scale. The negative relationship between judges' attitudes and item values reported by Hovland and Sherif (1952) is interpreted as a result of this difference in the scales adopted by in-range and out-of-range judges.

In his 1929 monograph Thurstone (Thurstone & Chave, 1929) staked the validity of the equal-appearing intervals model on his assumption that scale values are independent of judges' attitudes. On the basis of their evidence against this assumption, Hovland and Sherif (1952) argued that the validity of the model is in doubt. However, the present support for the variable series model suggests that judges' attitudes do not appreciably influence item values unless the item range is narrower than the range of judges' attitudes. From this perspective, it can be argued that Hovland and Sherif actually found only that Hinckley's (1932) item series was not valid for their judges. In theory it is possible to construct an item series which covers the positions of all members of any population of judges. If, in practice, a scale constructor cannot compile such a series, he appears to have the option of specifying a particular population of judges for whom the series can be expected to cover all own positions. This option would probably be satisfactory for most uses of attitude scales.

## SUMMARY

An experiment was performed to test four alternative interpretations of the influence of judges' attitudes on equal-appearing intervals scale values. The experimental design involved manipulation of the item series as well as judges' attitudes. Two of the theoretical models, perceptual vigilance and assimilation-contrast, can account for the obtained main effect due to judges' attitudes. Another of the models, adaptation level, can account for that effect, as well as the observed effect due to item series. Only one of the models, variable series, can account for both main effects and the significant interaction observed between judges' attitudes and item series.

Although the data generally support variable series, some of the obtained results do not. Notable, in this regard, is the finding of a positive relationship between own positions and scale values for extremely pro and moderately pro items.

On the basis of the general support for the variable series model, it was argued that valid equal-appearing intervals scales can be constructed only when the own positions of all judges are adequately covered by the item series. It was suggested that scale constructors specify the population of judges for whom this condition is met.

## REFERENCES

EDWARDS, A. L. A critique of "neutral" items in attitude scales constructed by the method of equal appearing intervals. *Psychol. Rev.*, 1946, **53**, 159–169.

FEHRER, ELIZABETH. Shifts in scale values of attitude statements as a function of the composition of the scale. *J. exp. Psychol.*, 1952, **44**, 179–188.

FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Ass.*, 1937, **32**, 675–701.

FRIEDMAN, M. A comparison of alternative tests of significance for the problem of *m* rankings. *Ann. math. Statist.*, 1940, **11**, 86–92.

HINCKLEY, E. D. The influence of individual opinion on construction of an attitude scale. *J. soc. Psychol.*, 1932, **3**, 283–296.

HOVLAND, C. I., HARVEY, O. J., & SHERIF, M. Assimilation and contrast effects in reactions to communication and attitude change. *J. abnorm. soc. Psychol.*, 1957, **55**, 244–252.

HOVLAND, C. I., & SHERIF, M. Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *J. abnorm. soc. Psychol.*, 1952, **47**, 822–832.

HUNT, W. A., & VOLKMANN, J. The anchoring of an affective scale. *Amer. J. Psychol.*, 1937, **49**, 88–92.

LIKERT, R. A technique for the measurement of attitudes. *Arch. Psychol., N. Y.*, 1932, No. 140.

MANIS, M. The interpretation of opinion statements as a function of recipient attitude. *J. abnorm. soc. Psychol.*, 1960, **60**, 340–344.

MICHELS, W. C., & HELSON, H. A reformulation of the Fechner law in terms of adaptation level applied to rating scale data. *Amer. J. Psychol.*, 1949, **62**, 355–368.

SEGALL, M. H. The effect of attitude and experience on judgments of controversial statements. *J. abnorm. soc. Psychol.*, 1959, **58**, 61–68.

THURSTONE, L. L., & CHAVE, E. J. *The measurement of attitude*. Chicago: Univer. Chicago Press, 1929.

VOLKMANN, J. Scales of judgment and their implications for social psychology. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads*. New York: Harper, 1951. Pp. 273–294.

WILCOXON, F. *Some rapid approximate statistical procedures*. New York: American Cyanamid, 1949.

(Received January 3, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 97–112

# THEORY AND EXPERIMENT ON THE MEASUREMENT OF APPROACH-AVOIDANCE CONFLICT[1]

SEYMOUR EPSTEIN AND WALTER D. FENZ

*University of Massachusetts*

A number of years ago, Luria (1932) introduced a technique for the measurement of conflict and reaction to stress. Among other procedures, he verified his approach by testing students before an unusually critical examination and by testing prisoners facing the death penalty and other serious sentences. The present study follows in the tradition of Luria's work in that an acute real life event is investigated, an attempt is made to measure affect selectively related to the event rather than the general affective state alone, and a specially devised word association test is used. The approach differs from Luria's in some of the measures investigated, in the substitution of a stimulus dimension for a dichotomous classification of the stimulus, and in the overall theoretical orientation. The critical event is a parachute jump for an inexperienced parachutist.

There is little question but that the first few jumps for a novice parachutist represent an acute approach-avoidance conflict between need for adventure, excitement, or prestige on the one hand, and fear of death and injury on the other. Expressions such as, "It was the most frightening experience of my life," or "For a while I thought I would not be able to go through with it," are commonly heard. Parachuting would seem to be comparable in momentary intensity to the events studied by Luria (1932), and clearly fulfills the conditions of an approach-avoidance conflict. In addition, parachuting has a unique advantage in that, unlike a murder or a critical examination, it can be repeated within short intervals, thereby allowing for the use of subjects as their own controls and for the investigation of experience and mastery. Finally, the frequency and timing of a parachute jump can be controlled according to the experimental plans of the investigator. In the present study, novice sport-parachutists were tested on the day of an anticipated jump and at a time 2 weeks from a jump. The general influence of parachuting and of time of testing upon responses to a stimulus dimension are investigated in relation to a theoretical model.

[1] This study is part of a project on the measurement of drive and conflict which is being supported by Grant M-1293 from the National Institute of Mental Health, United States Public Health Service. The study was also partially supported by a "teacher's research grant" from the University of Massachusetts.

Epstein is responsible for the theoretical portions of the paper. Fenz, who himself is a parachutist, conducted the study under the supervision of Epstein.

## *A Model for the Measurement of Conflict with Projective Techniques*

The model required is one which represents approach as stronger at the goal than avoidance, since subjects by their own choice advance to the goal. Such a model adapted from Miller (1948), is presented in Figure 1. It is assumed that there is an approach drive and an avoidance drive associated with parachute jumping, and that the approach drive can be represented by a less steep gradient as a function of an increasing dimension of stimulus relevance than the avoidance drive. It is assumed that drives have activating and directing properties which can be separately represented. It is assumed that the magnitude of drive produced increment in approach response can be represented by the algebraic difference of the approach and avoidance drives (see Figure 2). It is assumed that approach can be measured by goal relevant verbal responses and avoidance by the failure to produce such responses when they are normally elicited by the stimulus. It should be noted that the model permits the direct prediction of drive produced *increment* in approach response only and not overall drive relevance of the response, which is predominantly a function of the stimulus.

In regard to activation, it is assumed that

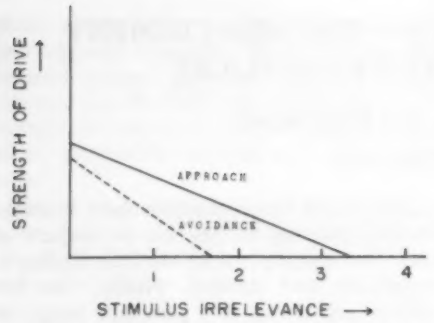



FIG. 1. Strength of approach and avoidance drive as a function of a stimulus dimension.

the magnitude of conflict produced activation can be represented by the sum of the magnitudes of the approach and avoidance drives disregarding algebraic sign (see Figure 3). It is assumed that activation can be measured by the GSR (Woodworth & Schlosberg, 1954), or by some other measure of autonomic activity, or a measure of somatic tension.

One advantage in having the gradients represent drives rather than responses or response tendencies, as in Miller's (1948, 1951) model, is that it immediately suggests the gradients be added if activation is to be considered, and allows for the prediction of two types of event from one basic model. A further reason for assuming a central concept of drive with activating and directive components is that naturalistic observation reveals that a person or animal brought into increasingly close contact (as measured by distance, goal relevant cues, or perceived availability) to a goal object exhibits increasing activation as well as stronger approach (or avoidance) responses. It is the unusual child who does not become perceptibly more excited (activated) as a wished-for event (e.g., Christmas morning) approaches, in addition to making approach responses (looking for and questioning about presents). Finally, not only is a central concept of drive helpful in referring to simultaneous changes in activation and goal direction brought about by the presence of a goal object, but, more important, the two aspects of drive are often reducible to each other. In the absence of experimentation it will be necessary once more to turn to observation and example.

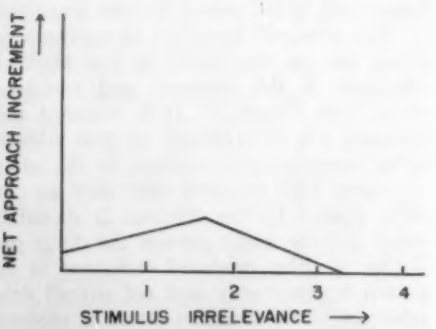



FIG. 2. Net approach increment as a function of a stimulus dimension.

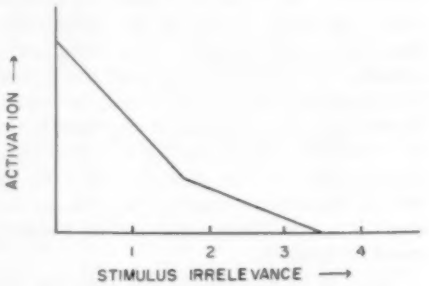



FIG. 3. Activation as a function of a stimulus dimension.

A dog belonging to one of the authors provides an excellent illustration. The animal makes intense approach responses in the form of running and leaping toward a ball, and clearly exhibits a gradient of approach by reacting more vigorously when the ball is near than when it is far. However, if he is commanded to sit still, and the ball made to approach the dog, a gradient of activation is observed in place of a gradient of approach. As the ball is brought closer, the dog becomes perceptibly more tense, his muscles appear to bulge through his skin, and a general tremor appears. We have no doubt but that if physiological measures of activation were recorded a steep gradient of activation would be indicated. Finally, when the ball is almost touching the animal's nose, the tension becomes unbearable and the dog barks vigorously, snaps at the ball, or entirely breaks the inhibition and dashes about the room.

While it is true that Miller allows for gradients of activation by relating avoidance to learned fear, his approach is different from our own in that we assume that activation is

a characteristic of all drives. We hypothesize that a gradient of activation can be demonstrated by using a barrier or physical restraint to prevent approach, but that gradients resulting from inhibition of approach responses are steeper, since they are determined by activation from two drives.

In adapting Miller's model to projective techniques, it is necessary to substitute verbal expression for approach and verbal inhibition for avoidance.[2] This raises the question of whether the principles of conflict that have been worked out by Miller and his colleagues in simple situations can, in fact, be translated into this different and more complex situation. The most critical assumption in the conflict and displacement models is that the gradient of avoidance is steeper than the gradient of approach. Miller (1948, 1951) indicates that this is not always the case but is contingent upon approach being more dependent than avoidance upon a physiological state. He indicates that the conditions are satisfied when approach is based upon hunger, and avoidance upon learned fear. If this is true, it would follow that inhibitory reactions should produce steeper gradients than the drives they inhibit, as the former are contingent upon the latter, and therefore more removed from any primary drive.

Assuming that the gradient of inhibition is steeper than the gradient of expression, the following hypotheses are indicated for the measurement of conflict with stimulus dimensions built into projective techniques:

1. Conflict is indicated by a sharp rise in activation as a function of increasing stimulus relevance.

2. Conflict is indicated by a relative increase in strength of approach responses to stimuli of low relevance and a relative decrease in strength of approach to stimuli of high relevance.

3. Conflict, when of sufficient magnitude, is indicated by a decrease in adequacy of performance as a function of increasing stimulus relevance. This last hypothesis follows from the consideration that high levels of activation are disruptive and that the directive effects of strong expressive and inhibitory drives result in inappropriate overemphasis upon certain stimuli and avoidance of others. The qualification that the conflict must be "of sufficient magnitude" is added to take into consideration that the curve of performance as a function of activation is probably inverted **U** shaped (Malmo, 1959).

### *Model Applied to the Present Study*

In the present study, parachutists were tested on the day of a jump and at a time 2 weeks removed from the jump. The model must, therefore, take into account time to a critical event. As a first step, models corresponding to Figures 1, 2, and 3, but with a time dimension substituted for a stimulus dimension, can be considered. If anticipated time to an event is taken as a measure of distance from the event, the proposed model corresponds to Miller's model for conflict. As the study simultaneously investigates the effects of two dimensions, it is necessary to consider a three-dimensional model. Such a model, which is analagous to combining Miller's models for conflict and displacement, is illustrated in Figure 4.[3] The curve for activation as a function of the stimulus dimension is taken from Figure 3 and reproduced on the z axis. A corresponding curve for activation as a function of a time dimension is represented on the x axis. Four points are selected along the time dimension, and the relationship of level of activation to the stim-

[2] It has come to our attention, after writing this paper, that Miller (1959), in his more recent work, has also related his model to projective techniques. However, he neither recommends the construction of stimulus dimensions nor the use of measures of activation. He believes that a projective response can be viewed as a displaced response, but that the relationship is complicated by additional considerations. (For an earlier discussion of the construction of stimulus dimensions in projective techniques to measure drive and conflict, see Epstein & Smith, 1956.) We agree with Miller on the complexity of the projective response, and, as will be seen later, find that predicting net approach increment as a function of conflict is theoretically complex in itself, apart from the consideration of other factors. Nevertheless, we believe that the difficulties may not be insurmountable, and even if they are for net approach increment, they will not be for measures of activation.

[3] See Murray and Berkun (1955) for an ingenious study in which a three-dimensional model of conflict and displacement is used to predict performance of rats in a runway with displacement alleys.

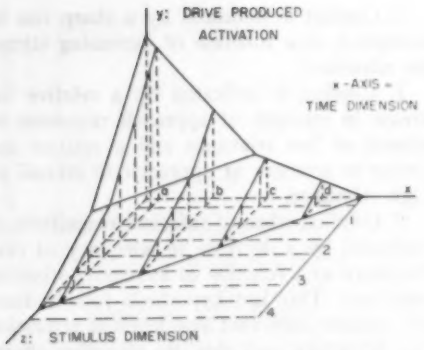



FIG. 4. Activation as a function of a stimulus and a time dimension.

ulus dimension for these points in time is indicated along the surface of the figure. The four resultant curves are extracted and represented to scale in Figure 5, which presents a family of curves for different time intervals to the jump. It is apparent that the gradients become higher and steeper as time to the jump increases.[4]

From Figure 5 and assumptions already made, the following predictions are indicated for the present experiment:

1. Parachutists at both times of testing will produce positive gradients as a function of an increasing stimulus dimension on the following measures: GSR, reaction time, and appropriateness of response. The prediction for reaction time follows from the consideration that for the word association test, response latency is an index of adequacy of response.

2. Parachutists tested on the day of an anticipated jump will produce higher and steeper gradients on each of the above measures than parachutists tested 2 weeks from a jump.

In order to predict the increment in approach response as a simultaneous function of a stimulus and time dimension, it is again necessary to consider a three-dimensional model. The curve for net approach increment from Figure 2 is reproduced on the x and z axes of Figure 6, with the x axis representing the time dimension and the z axis the stimulus dimension. The solid figure produced differs from the one for activation in that instead of a change in steepness there is a change in direction of slope. The surface can be described as a triangle leaning into a corner with the apex bent down to meet the intersection of the vertical walls at a point below the point of inflection. Four planes parallel to the yz plane represent the same four time intervals as in the figure for activation (see Figure 4). The intersection of these planes with the surface of the figure represents the relationship of net approach increment to the stimulus dimension for the time interval in question. The curves are extracted from Figure 6 and presented as a family of curves in Figure 7. Referring to Figures 6 and 7, it can be observed that the change in direction of the slope for both the time dimension and the stimulus dimension makes the prediction of net approach increment much more complex than was the case for activation. Depending upon the point along the time dimension selected, an increase along the stimulus dimension can result in an increase or a decrease in net approach increment. On either side of the point of inflection the same values occur, so that a lack of precision in locating the point of inflection is more critical than was the case for activation. Despite these difficulties, which prevent more specific predictions, the following would seem to be indi-

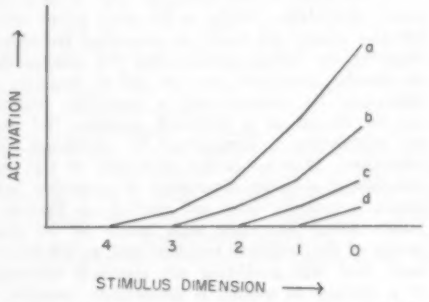



FIG. 5. Activation as a function of a stimulus dimension with time treated as a parameter. (The letters correspond to the letters in Figure 4, with a representing the shortest, and d the longest time interval.)

[4] It should be noted that the change in steepness of the slopes is a consequence of adding the two gradients. If either gradient alone were represented, a series of parallel curves differing only in height would be obtained. This may be observed in Curves c and d, which do not contain a component of the avoidance drive.

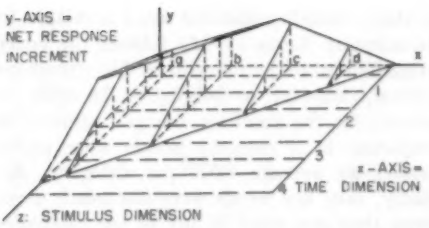



FIG. 6. Net approach increment as a function of a stimulus and a time dimension.

cated for the present study: (*a*) Parachutists at both times of testing will produce more parachute relevant responses than nonparachutists to words at the lower end of the stimulus dimension. (*b*) Parachutists on the day of an anticipated jump will produce more parachute relevant responses to stimuli of low parachute relevance and less parachute relevant responses to stimuli of high parachute relevance than parachutists tested 2 weeks from a jump, i.e., there is a significant interaction between time to a jump and stimulus relevance.

## METHOD

### *Subjects*

Experimental subjects were 16 novice parachutists from the University of Massachusetts Sport Parachuting Club. All had made at least one and not more than three parachute jumps at the time of testing. The subjects were recruited at the regular meeting of the parachute club and were asked to take part in the experiment for the payment of the costs of one parachute jump. The experimental group was matched with a comparable group of control subjects selected from students in an elementary course in psychology at the University of Massachusetts who volunteered to take part in experiments. Three additional subjects, who had considerable experience with parachuting, were also tested.

### *Materials*

Parallel forms of word association tests were constructed containing words with four levels of relevance to parachuting. The words were selected from a larger group rated on a four-point scale of parachute relevance by three experienced parachutists. Only words on which full agreement was reached were retained. The words were presented in increasing order of relevance to prevent highly relevant words from influencing responses to words of lower relevance. Three anxiety words were added to each list mainly out of curiosity, and were placed at the end of the list in order not to interfere with the dimension. One or more neutral words was presented between parachute relevant words and between anxiety words, and 10 neutral words were presented at the beginning of the list to reduce practice effects. Following are the word lists, with level of parachute relevance designated by figures, except for neutral words which were treated as the lowest level of relevance, and with anxiety words designated by the letter a:

*Word List A:* SALT, QUIET, RED, JOY, FOOTBALL, CITIZEN, MUSIC, CHAIR, SWISS CHEESE, HUNGRY, SKY$_1$, RHINOCEROS, LAND$_1$, BOOK, CATTLE, MOUNTAIN$_1$, MOON, WHITE, FALL$_2$, LOUD, DOG, FLOAT$_2$, SCISSORS, ALTITUDE$_2$, LION, SWIFT, BAIL-OUT$_2$, STOVE, RADIATOR, CANOPY$_2$, TAXI, GREEN, TARGET$_2$, SOUR, FEAR$_a$, BEAUTIFUL, BLACKOUT$_a$, MAN, INJURY$_a$, CARPET, HAPPY.

*Word List B:* MONKEY, MOON, RED, ELEPHANT, KEY, KING, COMFORT, SLEEP, SCREWDRIVER, CHICKEN, WIND, GREEN, SAND$_1$, PAPER, PENCIL, HILL$_1$, BLACK, STAR, DESCEND$_2$, HEAVY, FOOTBALL, OPENED$_2$, HAMMER, HARNESS$_2$, CAT, HARD, JUMP$_2$, LAMP, EAT, PARACHUTE$_2$, TOWN, PINK, RIPCORD$_2$, BITTER, ANXIETY$_a$, SOFT, KILLED$_a$, BOY, HURT$_a$, TABLE, FUN.

Words were presented by a tape recorder at 30-second intervals after having been screened by four judges for clarity of pronunciation. Subjects were instructed to respond as quickly as possible with the first word that occurred to them. Response words were recorded by the experimenter. Recordings of skin resistance were obtained from a direct current Hunter GSR apparatus, operating on the principle of a Wheatstone bridge. The apparatus was adapted for finger electrodes, and resistance was registered by an automatic recording pen. Time from end of stimulus to beginning of response was recorded by a pen-marker controlled by the experimenter.

### *Procedure*

Eight experimental subjects were tested on the day of a jump, and again 2 weeks after the jump; eight others were tested 2 weeks before a scheduled jump, and again on the day of the jump. Subjects were required to refrain from jumping for 2 weeks preceding the "before jump" testing and for 2

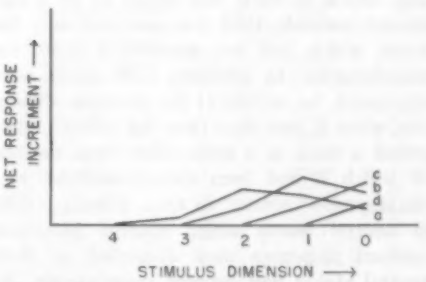



FIG. 7. Net approach increment as a function of a stimulus dimension with time treated as a parameter. (The letters correspond to the letters in Figure 6, with a representing the shortest, and d the longest time interval.)

weeks following the "after jump" testing, and to make no additional jumps between the testing sessions. Testing on the day of a jump took place 1–3 hours preceding the jump. Sixteen control subjects were tested twice with a 2-week interval between testing. In all cases, List A was presented in the first session and List B in the second session.

Upon entering the experimental room and being seated facing the examiner, a brief explanation was given about the word association test, during which the subject was told that the purpose of the test was to measure emotions related to parachuting. Electrodes were then attached to the right hand index and middle fingers. After about 3 minutes, to permit the electrodes to polarize and the subject to become accustomed to them, the tape recorded list of words was introduced with the comment:

> Now we will begin a test of speed of association to words. After you have heard a word—and not before you have heard the whole word—say the first word that occurs to you as quickly as possible. If you are not sure of a word, respond to what you think it was. Please do not make comments or ask questions between words, but save them for the end of the test. Any Question?

Where distributions permitted, the data were evaluated by analyses of variance. In one analysis, experimental subjects were used as their own controls, and nonparachutists were omitted. This permitted an orthogonal arrangement of experimental condition and order. In a second analysis, control subjects were compared on first and second testing.

## RESULTS

### *Galvanic Skin Response*

*Response to the stimulus dimension.* Galvanic skin response was measured by change in conductance immediately following presentation of a stimulus word. In order to make the results independent of selective perception, which in itself was found to be a significant variable, GSR was analyzed only for words which had not received a score for misperception. In addition, GSR scores were eliminated for words in the stimulus dimension when it was clear that the subject interpreted a word in a sense other than the one on which it had been dimensionalized, e.g., the response "winter" to FALL. Finally, GSRs to neutral words which elicited parachute relevant responses were discarded, as their neutral status had become questionable. By eliminating the complicating effects of perception, we could reasonably assume a meaningful stimulus dimension for each of the experimental subjects, instead of a dimension of the probability that the word is related to parachuting. Luria (1932) follows the latter procedure when he categorizes his stimuli as critical, doubtful, or indifferent. In order to correct for the resultant unequal number of responses in a category scores were represented by average GSR per category. Actually, very few words were discarded, never more than one word in any of the categories other than neutral, for which two was the maximum. The results would be essentially the same, but more difficult to interpret, had all words in a category been used. In representing the neutral words, the first 10 "practice" words were not counted.

Sixteen out of 16 experimental subjects produced a gradient of increasing GSR as a function of increasing stimulus relevance. One control subject also produced a gradient, but on questioning it turned out that he had been a paratrooper in the Army. His data were replaced by that of another control subject. For every experimental subject the gradient on the day of a jump was higher and steeper than on a control day. Although statistical analysis might appear to be superfluous, analyses of variance were carried out to investigate the effects of sequence, order, individual differences, and to allow for a comparison with the analyses of reaction time and of anxiety words. In addition, since the findings on the gradients were unequivocal, an opportunity was afforded for testing the relative validity of three commonly used indices of GSR.[5]

[5] The three measures were change in resistance as measured in ohms, change in conductance as measured in mhos, and change conductance as measured in log mhos. Results obtained from each measure were plotted for each experimental subject. On all three measures 16 of the 16 subjects showed a gradient of response corresponding to the dimension of stimulus relevance, and for each subject the gradient on the day of the jump was steeper than on the day remote from the jump.

Analysis of variance for each separate transformation yielded parallel results. A comparison of the three measures indicated that $F$ values for log conductance change were less pronounced than for resistance change and conductance change. This discrepancy may be attributed to the fact that the logarithmic measure reduced the distance between highly divergent scores. Logarithmic transformation is commonly used to reduce the relationship between conductance change and prestimulus level. Inspection

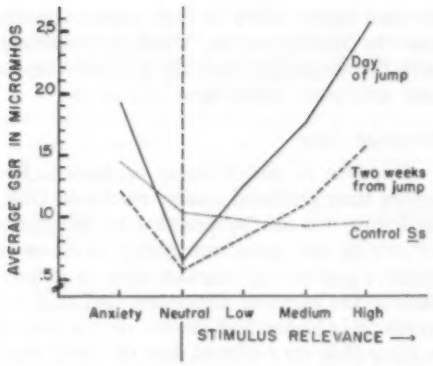



FIG. 8. GSR of parachutists and controls to a stimulus dimension and to anxiety words.

It is apparent that the control group produces no gradient (see Figure 8). The mean GSR in micromhos is 1.04 for neutral words and .95 for words of high relevance. The parachutists produce a gradient on the control day and a higher and steeper gradient on the day of a jump, all of which is in accordance with prediction.[6] The mean GSR in micromhos for neutral and high relevant words, respectively, is .56 and 1.57 on a control day, and .65 and 2.59 on the day of a jump.

Only the analysis of variance for parachutists serving as their own controls is presented, as no source of variance approached significance for the analysis of control subjects. In Table 1 it can be seen that there are significant differences for experimental condition, for stimulus dimension, and for the interaction of experimental condition and stimulus dimension. Significance is also indicated for the main effect of individual differences and for the interaction of individual differences with experimental condition. Apparently, not only do individuals differ reliably in the overall magnitude of their GSRs, but they also differ in the effect an approaching critical event has upon their GSRs. Perhaps this indicates an interesting personality variable that should be followed up in further work.

An additional finding of interest is that the parachutists have a smaller mean GSR than the control subjects to neutral words on both the day of a jump and the control day (see Figure 8). When tested by a separate analysis of variance for neutral words only, the difference is found to be significant at the .01 level.

TABLE 1

ANALYSIS OF VARIANCE OF GSR TO A STIMULUS DIMENSION FOR PARACHUTISTS ON THE DAY OF A JUMP AND A CONTROL DAY

| Source of variance | *df* | *MS* | *F* |
|---|---|---|---|
| Between-subjects | 15 | | |
| Sequence (Q) | 1 | 129.41 | 1.15 |
| Subjects | 14 | 111.70[a] | 11.67*** |
| Within-subjects | 112 | | |
| Group (G) | 1 | 795.51 | 17.07** |
| Q × G | 1 | 46.92 | 1.00 |
| Subjects × G | 14 | 46.58[a] | 4.87*** |
| Dimension (D) | 3 | 1309.91 | 95.12*** |
| Q × D | 3 | 11.01 | .79 |
| Subjects × D | 42 | 13.77[a] | 1.44 |
| G × D | 3 | 143.71 | 15.02*** |
| Q × G × D | 3 | 5.43 | .56 |
| Subjects × G × D | 42 | 9.57[a] | |
| Total | 127 | | |

[a] Error term for mean squares above it up to the next error term. The Subjects × G × D interaction was used to evaluate all other error terms.
** Significant at .01 level.
*** Significant at .001 level.

of the data indicated a positive relationship between prestimulus conductance level and magnitude of the deflection. In order to determine whether the results on GSR for the transformed measures of conductance could be attributed to differences in prestimulus level, the same analyses of variance were run on average prestimulus level as on GSR. The analysis of variance for prestimulus levels of experimental subjects on the day of the jump against prestimulus levels on the day remote from the jump did not yield a difference approaching significance. When experimental subjects under both conditions were compared to control subjects the analysis of prestimulus conductance measures also failed to yield significant values. The results for conductance change therefore could not be attributed to absolute conductance levels. That findings which were conclusively present by inspection were measured with least reliability by log conductance units throws some doubt on the advisability of this transformation for GSR data.

[6] The differential steepness of the gradients for the three groups cannot be attributed to the unequal heights of the gradients. The curves for the parachutists intersect the curve for the controls, and maintain their differential slopes with respect to each other within the same range on the ordinate. If anything, the difference in the heights of the gradients can be attributed to the difference in slopes, rather than the reverse, as the parachutists on the day of a jump and a control day do not differ significantly on neutral words.

TABLE 2

ANALYSIS OF VARIANCE OF GSR TO NEUTRAL AND ANXIETY WORDS FOR PARACHUTISTS ON THE DAY OF A JUMP AND A CONTROL DAY

| Source of variance | *df* | *MS* | *F* |
|---|---|---|---|
| Between-subjects | 15 | | |
| Sequence (Q) | 1 | 133.58 | 2.26 |
| Subjects | 14 | 59.00[a] | 6.88*** |
| Within-subjects | 48 | | |
| Group (G) | 1 | 267.62 | 4.46 |
| Q × G | 1 | 50.70 | .84 |
| Subjects × G | 14 | 59.94[a] | 6.97*** |
| Category (C) | 1 | 1483.10 | 29.17*** |
| Q × C | 1 | 72.58 | 1.42 |
| Subjects × C | 14 | 50.85[a] | 5.91** |
| G × C | 1 | 165.53 | 19.27*** |
| Q × G × C | 1 | 9.61 | 1.12 |
| Subjects × G × C | 14 | 8.59[a] | |
| Total | 63 | | |

[a] Error term for mean squares above it up to the next error term. The Subjects × G × C interaction was used to evaluate all other error terms.
** Significant at .01 level.
*** Significant at .001 level.

*Reaction to anxiety words.* A comparison of mean GSR to neutral and anxiety words (see Figure 8) indicates that both experimental and control subjects react more strongly to anxiety words than to neutral words. This difference is greatest for experimental subjects on the day of a jump, with a mean difference in micromhos of 1.28, next for experimental subjects on the day remote from a jump, with a mean difference of .64, and least for control subjects, with a mean difference of .41. As with neutral words the parachute group on a control day has a lower GSR to anxiety words than the control group. The analysis of the parachutists as their own controls is presented in Table 2, where it can be seen that not only do parachutists react significantly differently to anxiety and neutral words (C, .001 level) and that their differential reaction varies as a function of proximity to the jump (G × C, .001 level), but there are also significant individual differences in GSR across conditions (subjects), in change in GSR from control day to jump-day (Subjects × G), and in difference in GSR between neutral and anxiety words (Subjects × C). In the analysis of control subjects by themselves the difference between anxiety and neutral words is significant at the .05 level.

It is noteworthy that the parachutists on both the day of a jump and a control day produce larger GSRs to high relevant words than to anxiety words, which is consistent with the assumption that the gradients represent activation rather than fear alone.

*Reaction Time*

*Response to the stimulus dimension.* Reaction time produced similar results to GSR. In Figure 9 it can be seen that at both times of testing the parachute group produces a positive gradient of reaction time as a function of the stimulus dimension, and that the gradient is higher and steeper on the day of a jump than on a control day. On the control day, mean reaction time is 1.41″ for neutral words and 2.32″ for words of high parachute relevance. On the day of a jump the corresponding values are 1.45″ and 3.40″. The control subjects produce no gradient. Analysis of variance for the parachutists reveals that the difference associated with experimental condition (G), with the stimulus dimension (D), and with the interaction of experimental condition and stimulus dimension (G × D) are all significant (see Table 3). No source of variance is significant in the analysis of variance for the control group.

Of additional interest, and parallel to the findings in GSR, is the indication in Figure 9 that mean reaction time of parachutists to neutral words on a control day ($M = 1.41''$) is less than for control subjects ($M = 1.69''$), although in this case the analysis of variance of neutral words alone fails to reach significance.

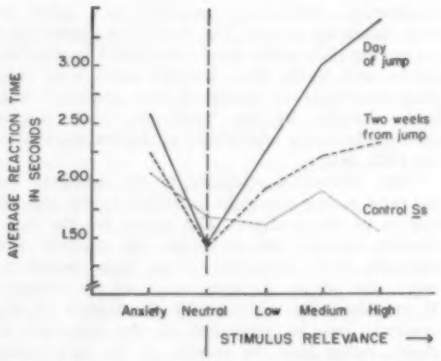



FIG. 9. Reaction time of parachutists and controls to a stimulus dimension and to anxiety words.

TABLE 3

ANALYSIS OF VARIANCE OF REACTION TIME TO A STIMULUS DIMENSION FOR PARACHUTISTS ON THE DAY OF A JUMP AND A CONTROL DAY

| Source of variance | *df* | *MS* | *F* |
|---|---|---|---|
| Between-subjects | 14 | | |
| Sequence (Q) | 1 | 1.34 | .42 |
| Subjects | 13 | 3.12 | 8.66*** |
| Within-subjects | 112 | | |
| Group (G) | 1 | 10.88 | 15.32** |
| Q × G | 1 | 2.13 | 3.00 |
| Subjects × G | 13 | .71[a] | 1.97* |
| Dimension (D) | 3 | 12.62 | 22.53*** |
| Q × D | 3 | .18 | .32 |
| Subjects × D | 39 | .56[a] | 1.55 |
| G × D | 3 | 1.62 | 4.50** |
| Q × G × D | 3 | .48 | 1.33 |
| Subjects × G × D | 39 | .36[a] | |
| Total | 119 | | |

[a] Error term for mean squares above it up to the next error term. The Subjects × G × D interaction was used to evaluate all other error terms.
* Significant at .05 level.
** Significant at .01 level.
*** Significant at .001 level.

TABLE 4

ANALYSIS OF VARIANCE OF REACTION TIME TO NEUTRAL AND ANXIETY WORDS FOR PARACHUTISTS ON THE DAY OF A JUMP AND A CONTROL DAY

| Source of variance | *df* | *MS* | *F* |
|---|---|---|---|
| Between-subjects | 14 | | |
| Sequence (Q) | 1 | .14 | .07 |
| Subjects | 13 | 2.07[a] | 2.46 |
| Within-subjects | 45 | | |
| Group (G) | 1 | 13.20 | 9.49** |
| Q × G | 1 | .73 | .86 |
| Subjects × G | 13 | 1.39[a] | 1.65 |
| Category (C) | 1 | 8.48 | 9.31** |
| Q × C | 1 | .64 | .76 |
| Subjects × C | 13 | .91[a] | 1.08 |
| G × C | 1 | 4.08 | 4.85* |
| Q × G × C | 1 | .31 | .36 |
| Subjects × G × C | 13 | .84[a] | |
| Total | 59 | | |

[a] Error term for mean squares above it up to the next error term. The Subjects × G × C interaction was used to evaluate all other error terms.
* Significant at .05 level.
** Significant at .01 level.

*Reaction to anxiety words.* A comparison of mean reaction time to neutral and anxiety words reveals that it took all groups longer to respond to anxiety than to neutral words (see Figure 9). The difference is greatest for parachutists on the day of a jump (1.12″), next for parachutists on a control day (0.84″), and least for control subjects (0.37″). The results are similar to those obtained for GSR (compare Figures 8 and 9). The findings are confirmed by analysis of variance. The analysis of parachutists on the day of a jump and a control day yields significance at the .01 level for experimental condition and for stimulus category, and significance at the .05 level for the interaction of experimental condition with stimulus category (see Table 4). The analysis of control subjects by themselves yields significance at the .05 level for the difference between neutral and anxiety words.

### *Failures in Perception*

*Perception of stimulus dimension and anxiety words.* Inspection of the data made it apparent that a considerable number of misperceptions and very few other obviously poor responses occurred. Failures in perception were scored when the subject reported that he had not heard a word, or when a response was given that was more appropriate to a word that sounded like the stimulus word than to the stimulus itself, such as the response "person" to SKY, where later questioning confirmed that the subject had heard SKY as "guy." Almost all inferred misperceptions were ones that occurred several times, and had been confirmed at least once by later questioning.

Figure 10 presents number of words missed over total words in a category, expressed as a percentage. Low and moderate parachute relevant words were combined to increase

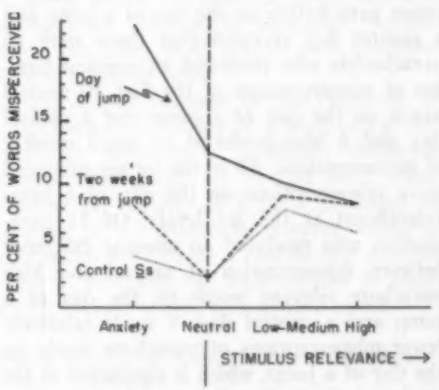



FIG. 10. Percentage of misperceptions by parachutists and control subjects of words on a stimulus dimension and anxiety words.

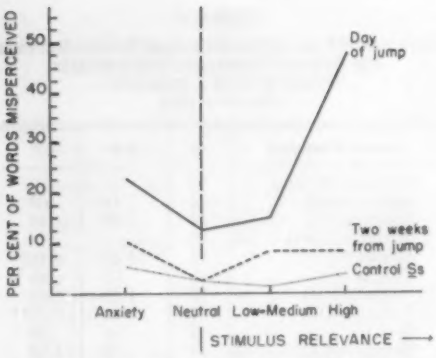



FIG. 11. Percentage of misperceptions of parachutists and control subjects of neutral words following words on a stimulus dimension and anxiety words.

reliability. Neutral words were restricted to the first 10 practice words because inspection (later verified) suggested that misperception of neutral words was influenced by the preceding parachute relevant word. It is evident in Figure 10 that parachutists on the day of a jump manifest a general perceptual deficit, as indicated by misperception of neutral words. There appears to be a specific deficit for anxiety words, and a relative sensitization [7] to parachute relevant words in relation to degree of relevance.

The high number of cases with zero misperceptions produced highly skewed distributions and accordingly chi square analysis of frequency data was used. A comparison between parachutists on the day of a jump and a control day revealed that there were 10 parachutists who produced an unequal number of misperceptions of the first 10 neutral words on the day of a jump and a control day and 6 who produced an equal number of misperceptions. All of the former produced more misperceptions on the day of a jump (significant at the .01 level). Of 11 parachutists who produced an unequal difference between misperception of neutral and high parachute relevant words on the day of a jump and a control day, 9 made relatively fewer misperceptions of parachute words on the day of a jump, which is significant at the .05 level. In addition, 12 out of 16 control subjects as compared to 6 out of 16 parachutists on the day of a jump produced a higher number of misperceptions of high parachute than neutral words, which is significant at the .05 level. Of 10 parachutists who misperceived an unequal number of anxiety words on the day of a jump and a control day, 9 made more misperceptions on the day of a jump, which is significant at the .02 level. Moreover, of 14 parachutists who produced an unequal difference in misperception of neutral and anxiety words on the day of a jump and a control day, 11 misperceived relatively more anxiety words on the day of a jump, which is significant at the .05 level. When comparison was with control subjects, 2 out of 16 controls and 11 out of 16 parachutists on the day of a jump misperceived more anxiety than neutral words, which is significant at the .01 level. Thus, on the basis of statistical analysis it may be concluded that parachutists on the day of a jump manifest a general perceptual deficit, as indicated in their misperception of neutral words, a specific deficit for anxiety words, and relative perceptual sensitization for parachute words.[8]

*Perception of neutral words following relevant words.* It has been noted that although

[7] Perceptual deficit and sensitization are used here descriptively only, and are not meant to imply a reason for the phenomenon.

[8] The findings cannot be explained away by inter- or intralist set effects. For one thing, if there were such set effects, they should influence the control subjects, yet there is no evidence of any perceptual effect for the control subjects. Second, parachutists served as their own controls, with half tested first on the day of a jump, and half tested first on a control day, and a careful analysis indicated that both subgroups behaved similarly. So far as intraserial set effects are concerned, the neutral words were presented before the parachute words, so that influence on the former was not possible. It might be argued that intraserial set effects could account for the sensitization to parachute words and the relative deficit for anxiety words, which appeared last. However, this would influence control subjects as well as parachutists on a control day, and cannot be a factor unless one is willing to postulate differential sets to pick up sets on the day of a jump, which is not too far off from selective perception. To determine whether such was the case, we examined the responses to successive words in the same category. There was no evidence for differential group performance as a function of an emerging set. It may be concluded that the data cannot be accounted for, directly or indirectly, by intra- or interlist set effects.

parachutists on the day of a jump had no difficulty in perceiving parachute related words they tended to misperceive the words that immediately followed. Accordingly, perception of neutral words was investigated as a function of the words preceding them. Figure 11 presents percentage of failures in perception for each of the categories of neutral words. It is evident that parachutists on the day of a jump produce a positive gradient of failures in perception as a function of an increasing stimulus dimension. There is also a tendency for parachutists on the day of a jump to misperceive neutral words following anxiety words.

Of 12 subjects who misperceived an unequal number of neutral words following words of high parachute relevance on the day of a jump and a control day, 10 had more misperceptions on the day of a jump, which is significant at the .05 level. When comparison is between control subjects and parachutists, 4 out of 16 controls and 10 out of 16 parachutists on the day of a jump misperceive at least one neutral word following a word of high parachute relevance, which is again significant at the .05 level. A comparison of misperception of neutral words following words of high relevance, relative to misperception of first 10 neutral words, reveals that of 15 parachutists who produced an unequal difference on the day of a jump and a control day, 12 made relatively more misperceptions of neutral words following parachute words on the day of a jump, which is significant at the .05 level. When the comparison is between parachutists and control subjects, 2 out of 16 controls and 9 out of 16 parachutists on the day of a jump misperceive relatively more neutral words following words of high parachute relevance than words among the first 10 neutral words, which is significant at the .01 level. It may be concluded that on the day of a jump there is a significant increase in the misperception of neutral words following words of high parachute relevance.[9]

[9] The possibility of intra- and interlist set effects producing the results was considered in the same manner as for the words in the dimension. Again, the results were negative. Luria (1932) also reports cognitive disturbance on neutral words following critical words.

Analysis of neutral words following anxiety words revealed that of eight parachutists who misperceived an unequal number of these words on the day of a jump and a control day, seven made more misperceptions on the day of a jump, which is significant at the .05 level. A comparison of misperception of neutral words following anxiety words relative to misperception of first 10 neutral words for parachutists on the day of a jump and either parachutists on a control day or control subjects approached, but did not reach significance ($p = .05–.10$). It may be concluded that there is an unreliable tendency for parachutists on the day of a jump to manifest perceptual deficit for neutral words following anxiety words.

### *Content of Associations*

It was predicted that parachutists on the day of a jump relative to a control day would demonstrate an increase in parachute relevant responses to stimuli of low relevance and a decrease in parachute relevant responses to stimuli of high relevance. Unfortunately, the prediction could not be tested at the high end of the dimension as stimuli of high relevance almost always elicited responses that were relevant to parachuting, and we were unable to devise a reliable scoring system for determining degree of approach within this range. As for the rest of the dimension, there was no evidence that responses to neutral words and to words of low and moderate parachute relevance were functionally dissimilar. Accordingly, responses to all stimuli were combined into one overall score.

Analysis of variance of the responses of parachutists on the day of a jump and a control day revealed a significant difference between the groups at the .01 level. Parachutists on the day of a jump produced an average of 8.3 parachute relevant words as contrasted with 5.8 on a control day. Control subjects produced the least number of parachute relevant words, with a mean of 3.2.

### *Three Records of Experienced Parachutists*

Although the reliability of results based on a sample as small as three is necessarily low, the results were so striking that we decided

they were worth reporting. For both mean GSR and mean reaction time, the gradients for the experienced and inexperienced parachutists were alike on the control day. However, while the inexperienced jumpers produced a monotonic curve on the day of a jump, the experienced jumpers produced a curve similar to it up to an intermediate point on the stimulus dimension, after which it dropped off so sharply that the mean reaction at the upper end of the dimension was no different from what it was at the lower end of the dimension. All 3 experienced jumpers and none of the 16 inexperienced jumpers or 16 controls produced this curve.[10]

## DISCUSSION

### *Measures of Activation*

It had been assumed that GSR is a direct measure of activation and that reaction time in a word association test is a measure of adequacy of performance, and therefore, indirectly of activation. Accordingly it was predicted that parachutists would produce similar gradients for both measures on a control day, and that the gradients would be higher and steeper on the day of a jump. Apart from supporting the model, the high reliability of the findings with these two measures indicates that they are completely adequate for measuring individual differences. Considering that reaction time and particularly GSR are indirect measures of conflict and not normally under conscious control of the individual, they are particularly promising measures of unconscious conflict. The results are striking enough to warrant testing of known clinical groups with specially devised stimulus dimensions used in conjunction with measures of GSR and reaction time. As an example, ulcer patients might be tested with stimulus dimensions of hostility and dependency built into a word association test or a TAT, and schizophrenics might be tested with stimulus dimensions of nurturance and rejection. A study along the latter lines is presently under way. The approach need not be confined to broad motive states, but specific sources of conflict can be explored as well.

The finding that reaction time produced results almost identical to GSR is consistent with Lanier's (1941) conclusion that the affective change brought about by words representing a conflict influences length of reaction time in the same direction and to about the same extent as GSR. However, it is evident that the relationship between GSR and reaction time is not a direct and simple one. Hathaway (1929) reported a correlation of .60 between GSR and reaction time, while Hunt and Landis (1935) found no relationship. Peterson and Jung (1907) reported that only in certain cases is there a clear relationship. What is probably the case is that GSR and reaction time measure different aspects of conflict. Following the model presented, GSR is a measure of activation, while reaction time, at least in a word association test, is a measure of adequacy of performance, which is a frequent but not necessary concomitant of activation. Assuming that the curve of level of performance as a function of increasing activation is inverted **U** shaped, the relationship between reaction time and activation would depend in part upon the nature of the task, and in part upon the degree and range of activation sampled. Moreover, in a test where the subject was not required to respond rapidly, reaction time would not be a measure of adequacy of performance. Finally, although increases in activation within the above limitations should increase reaction time, the reverse need not hold and reaction time could be affected by many sources that have no bearing on GSR. To the extent that both measures in a word association test are more highly correlated with conflict than with each other, they should provide useful supplementary measures of conflict.

### *Measures of Approach and Avoidance*

*Perceptual sensitization and defense.* It had been predicted that positive gradients would be produced by measures of misperception and reaction time on the assumption that both were measures of adequacy of response. Although the prediction was substantiated for reaction time, a significant opposite relationship was found for misperception. Yet, the

[10] We have now, as part of another study, confirmed the results with three more experienced parachutists.

basic assumptions from which the prediction for misperceptions was made were supported, i.e., the findings on GSR provided evidence of an increase in activation along the stimulus dimension, and the misperception of neutral words following words on the dimension provided evidence that activation interferes with perception. What had not been taken into account is that perception of words along a dimension is a measure of selective content in much the same maner that content of association is. In the present study approach was greater than avoidance, so that net approach increment could be expected to compensate for interference effects from activation. It is as if the subject, despite a general state of high tension and further increases in tension brought about by thoughts of parachuting, nevertheless forces himself to concentrate on parachuting. The effect is to overcome anxiety induced perceptual deficit for parachute relevant stimuli. The adaptive significance of such a reaction is apparent. It is not possible to explain away the findings by word frequencies, familiarity with words, or inter- or intraserial set effects. Thus, the evidence would appear to support unequivocally the phenomenon of perceptual sensitization. The findings on anxiety words equally well support selective perceptual deficit, which in light of the experimental controls and negative findings on set effects, is strongly suggestive of perceptual defense, i.e., an anxiety reducing defense mechanism. On the day of a jump novice parachutists are in a state of acute anxiety, with which they must come to grips if it is not to interfere with their performance and endanger their lives (Walk, 1959 unpublished). Thinking about and being oriented toward anxiety producing stimuli, i.e., worrying, would only increase level of anxiety, so that there is a strong basis for an avoidance reaction to such thoughts and perceptions. In order to check the anxiety reducing effects of misperception further, we examined the GSRs to anxiety words that were misperceived. Where an innocuous word was perceived in place of an anxiety word, the GSR was reduced. Thus, the overall evidence suggests that perceptual defense did occur for the anxiety words. Although originally unanticipated, perceptual sensitization and defense may well provide effective measures of approach and avoidance. Of course, in situations where unconscious conflict was involved, avoidance would be stronger than approach, and net approach tendency would work in the same direction as activation, so that enhanced perceptual defense might be the only phenomenon that would be found.

It is noteworthy that failures in perception were far more common than all other inadequate responses combined. Luria (1932) noted failures in perception for emotionally disturbing words, but these were no more common than were other inadequate responses, such as perservation, repeating a stimulus word, making an incidental comment, or blocking. The larger number of misperceptions in the present study undoubtedly is a result of the tape recorded presentation, despite the screening of the lists for clarity of pronunciation. Tape recordings obviously provide reduced cues because of the absence of lip movements. It is possible that tape recorded lists containing specially selected categories of words would provide an excellent projective technique for general usage, particularly if fidelity or loudness were varied to provide an optimum level of ambiguity or a measure of perceptual threshold.

*Content of association.* The results on content of association corresponded to those on perceptual sensitization, i.e., parachutists on the day of a jump produced an increased number of parachute related responses. Unfortunately, the critical prediction so far as the model is concerned, that conflict would be indicated by an increment in parachute relevant responses at the low end of the parachute dimension and a decrement at the high end, could not be adequately tested because all associations to the high end of the dimension were uniformly high in relevance. To test the hypothesis it will be necessary to use words at the upper end of the stimulus dimension which facilitate avoidance reactions to a greater extent than the ones used in the present study. Nevertheless, the reduced number of misperceptions at the high end of the dimension leads one to suspect that the hypothesis would not be substantiated. Moreover, there was no evidence within the range of the dimension that could be adequately

tested indicating that responses to different levels of the dimension were functionally dissimilar.

The measures of perception and of content of association clearly illustrate the manner in which conflict serves to focus attention in the area of conflict and to produce insensitivity in other areas. The implications for psychopathology are evident, particularly in reference to symptoms involving preoccupation, such as obsessive thinking, and to symptoms of inefficiency and inability to concentrate.

*Effect of Experience on the Shape of Gradients*

The findings on three experienced parachutists are particularly interesting. A reasonable prediction would have been that experienced jumpers would produce gradients that were lower and less steep than inexperienced jumpers, particularly on the day of a jump, i.e., their gradients would be somewhere between an inexperienced jumper's and a nonparachutists'. However, it was found for both GSR and reaction time that on the day of a jump, experienced jumpers produced similar gradients to inexperienced jumpers up to a midway point along the dimension, after which their gradients reversed direction, so that their GSRs to high relevant parachute words were no different than to neutral words. The resulting GSR and reaction time curves for the experienced jumpers resembled the predicted curve for net approach increment, which was based on the assumption that inhibitory tendencies are in conflict with expressive tendencies. This may indicate that associated with experience and mastery of conflict there is an inhibition of anxiety producing responses. Another possibility, not incompatible with the first, is that the gradient of activation is steeper for cues that are consciously associated with the conflict than for cues that are indirectly conditioned to the conflict, and that with successful experience, the anxiety associated with the labeled cues is extinguished at a faster rate than the anxiety conditioned to the unlabeled cues. Parachutists, in general, are more apt to consciously relate conflict produced by parachuting to high relevant words, such as PARACHUTE and RIPCORD, than to low relevant words, such as LAND and HILL. It would then follow that with successful experience conflict produced fear should be extinguished faster to the high relevant words than to the low relevant words. If the argument here is correct, it offers an explanation of why anxiety is often elicited by apparently innocuous or unrecognized stimuli, and why it is therapeutic to label stimuli associated with a conflict. In reference to this latter point, an ex-paratrooper colleague told us of how he had experienced an anxiety attack while standing on the platform of a train, until it occurred to him that the perception of the terrain moving by was similar to his perceptions when preparing to jump from an airplane.

*Anxiety Reducing Effects of Parachuting*

Parachutists obtained lower GSRs and shorter reaction times to neutral stimuli than did control subjects. Possibly, this indicates a personality characteristic which leads people to choose parachuting as a sport. Short reaction times and small GSRs suggest a person who is relatively impulsive and expressive. A second possibility is that parachuting serves to reduce tensions, and that the focusing of conflict and its momentary mastery in one area temporarily reduces conflict and disturbance in other areas, i.e., parachuting is a way of mastering fear. A few records of parachutists tested immediately after a jump revealed an astonishing absence of GSRs. The father of one of the parachutists reported that his son, who is a chronic stutterer, experienced an alleviation of symptoms on the afternoon following his first jump. Another subject, in a letter to the examiner, volunteered information about his increased frustration tolerance following a parachute jump. Reference is here made to the letter, which provides a description of a first jump:

> Well, my first static line jump was an experience that left some definite impressions on me . . . . Yes, fear of the unknown was there. I had prepared for this a long time. It was a real challenge to me. I had fears of chickening out, fear that the chute might not open, or that my landing would be hard. I was very anxious to get it over with. The most immediate sequence that is constantly on my mind is the time interval of standing on the wheel and holding on to the struts, being hit by the jumpmaster, and falling backwards till my chute caught.

What happened after that is only slightly in my mind. After landing I was happy, emotionally released, and proud of myself in that I had accomplished something I set out to do. This was a good morale booster and reduced my insecurity . . . . While driving back we tried to pick up two girls on the road. They refused, but this did not shatter my ego. I had an "I don't care" attitude. While on the road I did not speed, or at least there was no desire to speed. Another change I noticed was that when I put on the taperecorder to hear some music, I did not want to hear it loud. That night I finished reading a book, but while reading, the experience of jumping was constantly on my mind. . . .

## SUMMARY

A theoretical model for the measurement of approach-avoidance conflict was described, and predictions from it were evaluated by examining inexperienced parachutists at two time intervals from a jump. The model is a modification of Miller's model for approach-avoidance conflict. Gradients of drive were represented as a function of a stimulus and a time dimension. From the basic model of drive strength, models of activation and net approach increment were derived. It was assumed that the former could be measured by GSR and by performance deficit, and the latter by verbal content in response to a word association test with a built-in stimulus dimension. It was predicted that parachutists would produce gradients for GSR and for measures of performance deficit, and that the gradients would be higher and steeper on the day of a jump than at a time 2 weeks from a jump. It was further predicted that parachutists on the day of a jump would produce an increased number of goal relevant responses to the lower end of the stimulus dimension and a decreased number of such responses to the upper end of the dimension relative to their performance 2 weeks from a jump.

Eight novice sport-parachutists were tested first on the day of a jump and again 2 weeks later; eight others were tested first 2 weeks before a jump and again on the day of a jump. Sixteen nonparachutists served as controls. The tests consisted of parallel forms of a word association test with four levels of relevance to parachuting and with a group of words related to anxiety.

The major findings were as follows:

1. All 16 parachutists on both the day of a jump and the control day, and no control subjects, produced gradients of GSR as a function of the stimulus dimension. For all 16 parachutists, the gradients were higher and steeper on the day of a jump than on a control day.

2. Anxiety words elicited an increase in GSR from controls as well as parachutists. The increase was greatest for parachutists on the day of a jump, next for parachutists 2 weeks from a jump, and least for controls.

3. Parachutists produced smaller GSRs than nonparachutists to neutral words.

4. The results on reaction time, a measure of performance deficit, were similar to those for GSR.

5. Parachutists on the day of a jump exhibited a general perceptual deficit for neutral words, a greater deficit for anxiety words, and a relative sensitivity for parachute related words. The same results were obtained when parachutists were used as their own controls as when they were compared to nonparachutists. The findings can neither be accounted for by general frequency of or familiarity with words, or by inter- or intraserial set effects.

6. Parachutists on the day of a jump produced more parachute related responses than parachutists on a control day, who in turn produced more such responses than nonparachutists.

7. Three experienced jumpers on the day of a jump all produced a different form of curve for GSR and reaction time than inexperienced jumpers. The curve was similar up to a midpoint in the dimension, after which there was a sharp drop for the experienced jumpers and a continuous increase for the inexperienced jumpers.

It was concluded that the following three effects appear to be useful indicators of conflict: an increase in activation along a stimulus dimension, as indicated by a physiological measure, such as GSR; an increase in performance deficit along a stimulus dimension, as indicated by a formal (noncontent) measure, such as reaction time; and selective approach and avoidance, as indicated by perceptual sensitization and perceptual deficit, or by content of verbal response.

## REFERENCES

EPSTEIN, S., & SMITH, R. Thematic apperception as a measure of the hunger drive. *J. proj. Tech.*, 1956, **20**, 372–384.

HATHAWAY, S. R. A comparative study of the psychogalvanic and association time measures: A new psychogalvanic apparatus. *J. appl. Psychol.*, 1929, **13**, 632–646.

HUNT, W. A., & LANDIS, C. Word association, reaction time, and the magnitude of the GSR. *Amer. J. Psychol.*, 1935, **47**, 143–145.

LANIER, L. H. An experimental study of "affective conflict." *J. Psychol.*, 1941, **11**, 199–217.

LURIA, R. *The nature of human conflicts*. New York: Liverright, 1932.

MALMO, R. B. Activation: A neuropsychological dimension. *Psychol. Rev.*, 1959, **66**, 367–386.

MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.

MILLER, N. E. Comments on theoretical models illustrated by the development of a theory of conflict behavior. *J. Pers.*, 1951, **20**, 82–100.

MILLER, N. E. Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A study of a science*. New York: McGraw-Hill, 1959. Pp. 196–292.

MURRAY, E. J., & BERKUN, M. M. Displacement as a function of conflict. *J. abnorm. soc. Psychol.*, 1955, **51**, 47–56.

PETERSON, F., & JUNG, C. G. Psychological investigations with the galvanometer and pneumograph in normal and insane individuals. *Brain*, 1907, **30**, 211.

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. New York: Holt, 1954.

(Received January 16, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 113–120

# MEDIATION AND ASSOCIATIVE FACILITATION IN NEUROTIC, PSYCHOTIC, AND NORMAL SUBJECTS[1]

PETER J. LANG AND KENNETH LUOTO

*University of Pittsburgh*

The mediation hypothesis assumes that the learning of a new response to a given stimulus may be facilitated if an association exists between this new response and some previously learned response to that stimulus. There is little general agreement as to the exact nature of the process (*rg*, MPS, cell assembly), but it seems impossible to account for complex behavior without a mediation construct. Osgood (1953) assumes mediation in his discussion of thinking. For Dollard and Miller (1950) verbal labels mediate emotional responses in psychotherapy. Goss (1958), Griffith, Spitz, and Lipman (1959) suggest that verbal mediators are fundamental to concept formation. For many, "the mediation postulate is made to carry the burden of insight, problem solving, and efficiency in transfer situations" (Bugelski, 1956).

Despite its importance to theories of thinking and verbal behavior, direct evidence for an implicit mediation process in verbal learning rests mainly with studies of college students (Bugelski & Scharlock, 1952; Kitao, 1960; Russell & Storms, 1955). The present research proposes to test the mediation hypothesis with an adult, noncollege sample of normal subjects, and to examine differences between these subjects and patients displaying behavior pathology. Two subgroups of the latter class, anxiety neurotics and schizophrenics, are of particular interest.

In a discussion of the differential performance of high and low anxious subjects in a paired-associate task, Spence (1958) suggests that the former will perform at a higher level if "the associative connections between each stimulus word and the nonpaired response words are lower than that to the paired word" (p. 139). These conditions are obtained in a list, such as the one employed in the present research, which is composed of both pairs assumed to be related by a mediating term and pairs not so selected. Thus, we might anticipate that anxious subjects would show greater relative benefit from the mediational connection than "normal" subjects.

However, a contrary prediction is also applicable to the present research. In this experiment, the subject learns two paired-associate lists. Only the response terms are altered in the second list. Thus, in List II, the correct response is at the outset weaker than a potential intruder. Given approximately these conditions, Spence (1958) suggests that "the stronger the drive, the greater will be the percent choice of the wrong response, or, in other words, the poorer will be their performance at this initial stage" (p. 138). Although this prediction is relevant to the entire list, the conditions are especially enhanced for the nonmediation pairs. The list was constructed in such a way that, for these pairs, each stimulus member may be more closely related to another response term in the list than the one the subject is required to learn as part of the pair. Thus, considerable interference might be anticipated, retarding particularly the performance of anxious subjects on this part of the list. The present research will help determine the relevance of these drive theory predictions to mediation processes in verbal learning.

Since Bleuler (1950) disturbance of association has been considered a primary symptom of schizophrenia. This disturbance has been variously interpreted as a disorder of conceptual ability (Hanfmann & Kasanin, 1942), abstract behavior (Goldstein, 1939), or "looseness of association" (Meadow, Greenblatt, & Solomon, 1953). All of these imply either a defect of the mediational process in psychosis, or at least, a fundamental difference in the manner of its functioning in the

[1] This investigation was supported by United States Public Health Service Research Grant M-3567, Associative Transfer and Interference in Schizophrenia.

normal and schizophrenic patient. It would, thus, seem important to determine if the mediational effect in paired-associate learning can be demonstrated in the latter population, as it has been with college students.

Furthermore, Mednick (1958) has suggested that the thinking disorder of schizophrenics is a function of current anxiety, or represents the residual of a chronic anxiety state. His predictions of schizophrenic behavior in the setting under consideration parallel those of Spence for high drive subjects, with the qualification that the effects might be more extreme for the psychotic.

## METHOD

### *Experimental Design*

The design of this experiment was similar to that employed in their studies of college subjects by Russell and Storms (1955) and Bugelski and Scharlock (1952). In the latter experiment, subjects learned three lists of paired associations: List I established A-B associations; List II, B-C associations; List III tested the mediation effect by comparing the learning of A-C with A-X pairs.

Russell and Storms (1955) simplified the design by eliminating the learning of List II. In their study the B terms of List I were all items taken from the Kent-Rosanoff Word Association Test. The response terms used in the test list were associates of the B terms as defined by the normative data of the test.

The present design represents a further improvement, necessitated by the employment of pathological populations. Subjects were first administered a part of the Kent-Rosanoff Word Association Test, which contained among the 33 items administered the 10 B terms to be used in this experiment. The unique associates of each subject were subsequently used as the response terms in the second list.

### *Construction of the Paired-Associate Lists*

The first list learned by all subjects was the 10-pair nonsense syllable word list used as a first list by Russell and Storms (1955). The second list was composed of the same stimulus terms used in List I. Half of these stimuli were paired with response words that were mediated associations; the other five pairs (neutral) were made up of unrelated stimulus and response terms. However, the specific word used as responses for both mediated and neutral pairs varied systematically from subject to subject. For the first experimental subject, the first five response terms were the subject's unique associates to the List I response terms (taken earlier from the Word Association Test). The subject's own associations were also used for the second five response terms. However, they were assigned to stimuli randomly, with the restriction that each stimulus was assigned a response term unrelated to its response of List I. List II for subsequent subjects was constructed in a similar manner, except that the five stimuli receiving mediated and the five receiving unrelated (neutral) response terms were alternated through the sample. Thus, differential learning of mediated and nonmediated pairs could not be ascribed to the ease of learning a particular set of nonsense syllables or the associates of any particular set of response terms.

### *Subjects*

The design described above was replicated with three different subject populations. The first group was composed of 15 patients convalescing on the surgical wards of the Oakland Veterans Administration Hospital, Pittsburgh, Pennsylvania. They presented no history of psychiatric disorder; patients with physical complaints frequently termed "psychosomatic"—stomach ulcers, low back pains, asthma—were also excluded from this group. Their mean age was 32.13 years. The average Wechsler Vocabulary score was 23.53, and the *MA* scale stanine was 4.73.

The second group consisted of 15 patients from neuropsychiatric wards of the Oakland Veterans Administration Hospital. All these subjects carried a diagnosis of "anxiety reaction" or "anxiety neurosis" and were judged to be anxious by attendant personnel. The mean *MA* scale stanine for this group was 8.2. They displayed no psychotic symptoms nor revealed any history of psychosis. The mean age of the anxious subjects was 35.13 years. Their vocabulary score was 26.06.

Fifteen chronic schizophrenic patients in partial remission made up the third group. They were seen at the Leech Farm Veterans Administration Hospital, Pittsburgh, Pennsylvania. These subjects had no history of neurological involvement or primary mental deficiency. Their vocabulary score was 29.73. The average age of this group was 30.10 years.

In general, all male patients between the ages of 20 and 41, who met the diagnostic criteria of the appropriate subgroup, had an equal opportunity to participate in the experiment. However, this condition was overridden by either of the following factors: refusal to participate; the subject judged to be "out of contact and untestable" after consultation between the experimenter and the ward attendant. Furthermore, if a patient had learned less than five pairs after 30 presentations of the entire list, it was held that the task was too difficult to be meaningful for this subject, and he was dropped from the group. Six subjects were eliminated for this reason.

### *Procedure*

The subject was first administered the Word Association Test, followed by the Wechsler-Bellevue Vocabulary list. Subsequently the subject learned two

TABLE 1

MEAN NUMBER OF TRIALS REQUIRED TO ACHIEVE CRITERION FOR MEDIATED PAIRS, NEUTRAL PAIRS, AND THE TOTAL LIST

| Diagnosis | List | Mediation | Neutral | Total |
|---|---|---|---|---|
| Normal | I | 14.31 | 14.53 | 14.42 |
| | II | 9.67 | 11.53 | 10.58 |
| Anxiety | I | 12.76 | 13.75 | 13.25 |
| | II | 9.25 | 12.35 | 10.75 |
| Schizophrenia | I | 14.39 | 13.95 | 14.17 |
| | II | 11.64 | 14.93 | 13.30 |

lists of nonsense syllable-word pairs, with a 15-minute interval between lists. The pairs were typed on individual 3″ × 5″ cards and viewed by the subject through an opening in a black, plywood panel. Behind the panel a motor driven, slotted screen was arranged to expose the stimulus half of each pair for 2 seconds and the response half for an additional 3 seconds, with 3.5 seconds between pairs (Bugelski & Scharlock, 1952).

The subject learned to anticipate each correct response to a criterion of four successive, perfect trials. As the criterion for each pair was met, that pair was dropped from the list to prevent overlearning. The cards were shuffled between trials to eliminate serial learning. The entire task was usually completed in less than an hour and a half, including the interval between lists during which the subject relaxed with a magazine. Following the experimental session, the neurotic and nonpsychiatric patients were administered a short form of the Taylor *MA* scale (Bendig, 1956).

All correct anticipations and errors were tabulated for each subject, for both Lists I and II. The measures used in the final analysis were: the mean number of trials required to reach criterion on all pairs in List I, and separately for List II; the mean number of trials required to reach criterion on all five mediated pairs, and separately on all five neutral pairs, for both Lists I and II. For List I "mediated pairs" refers to those stimuli which acquired the mediated response term in List II.

These measures were used because they seemed to provide the least biased estimate of overall performance. However, they offer no information about the rate or progress of learning. Therefore, the number of trials required to reach criterion on each of the 10 successive pairs in List I and II were also tabulated.

In discussing the relevance of drive theory to paired-associate learning, Spence (1958) makes the following point: ". . . the theory of paired-associates learning has as yet not been developed sufficiently to predict what will happen beyond the first few trials, and it would have been more appropriate, as far as implications for our drive theory are concerned, if we had used at most only the data from the first four or five trials" (p. 139). In accordance with this admonishment, the number of correct responses made in the first five trials of List II were separately recorded for the total list, the mediated pairs, and the neutral pairs.

## RESULTS

Despite differences in diagnosis, no difference in initial learning between subgroups was observed. A variance analysis of the mean number of trials required to learn List I resulted in an $F$ less than 1.0. Criterion scores achieved by each patient group on both Lists I and II are presented in Table 1.

The statistic used to evaluate all these data was the analysis of variance for repeated measures, in which there are two variables within-subjects (list and mediation) and one between-subjects (hospital diagnosis). Because specific hypotheses were raised concerning differences between psychotic populations and between anxious (high drive) and nonanxious (low drive) subjects, these com-

TABLE 2

ANALYSIS OF VARIANCE OF MEAN TRIALS TO CRITERION FOR DIAGNOSIS, LIST, AND MEDIATION

| Source | *df* | *MS* | *F* |
|---|---|---|---|
| Diagnosis | 2 | 46.04 | |
| Psychosis (Schizophrenic vs. Anxiety and Normal) | 1 | 85.08 | 1.10 |
| Anxiety (Anxiety vs. Normal) | 1 | 7.00 | — |
| Pooled subjects with ingroups | 42 | 77.33 | |
| Lists | 1 | 255.85 | 11.80** |
| Diagnosis × Lists | 2 | 32.47 | 1.49 |
| Psychosis × Lists | 1 | 50.93 | 2.34 |
| Anxiety × Lists | 1 | 14.01 | — |
| Pooled subjects × Lists within diagnosis | 42 | 21.68 | |
| Mediation | 1 | 101.85 | 11.86** |
| Diagnosis × Mediation | 2 | 3.77 | — |
| Psychosis × Mediation | 1 | .13 | — |
| Anxiety × Mediation | 1 | 7.41 | — |
| Pooled subjects × Mediation within diagnosis | 42 | 8.59 | |
| Lists × Mediation | 1 | 69.94 | 8.16* |
| Diagnosis × Lists × Mediation | 2 | 4.53 | — |
| Psychosis × Lists × Mediation | 1 | 8.64 | 1.008 |
| Anxiety × Lists × Mediation | 1 | .41 | — |
| Pooled subjects × Lists × Mediation within diagnosis | 42 | 8.57 | |

Note.—The between diagnosis variance has been further subdivided into two components called Psychosis, the schizophrenic group vs. the anxiety and normal groups; and Anxiety, the anxiety group vs. the normal group.

* $p < .02$.
** $p < .01$.

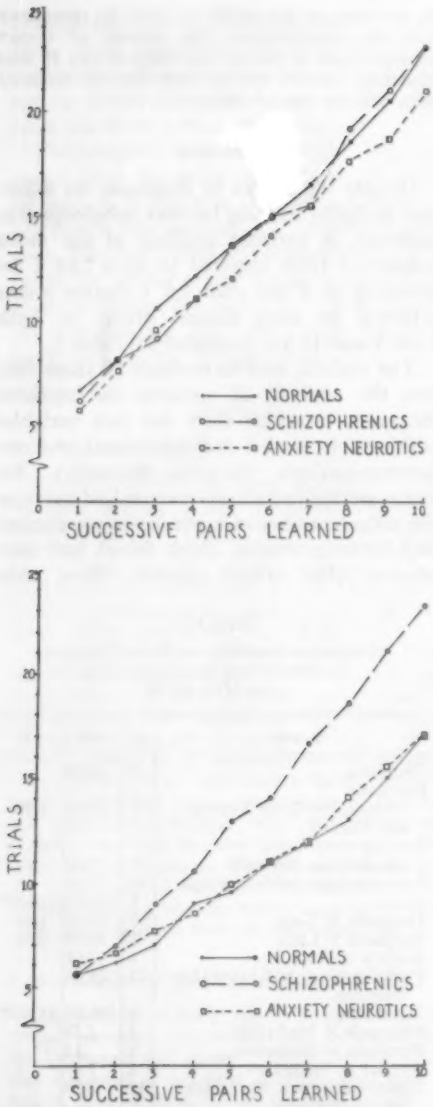



FIG. 1. Mean number of trials required to learn successive pairs for List I and List II.

ponents of the between-groups variance were separately analyzed.[2]

The results of this analysis are presented in Table 2. The Lists × Mediation $F$ value is significant beyond the .02 level, providing clearcut evidence that the mediation process extends across diagnostic groups. Furthermore, there is no indication in this analysis of a differential mediation effect for either anxious or psychotic subjects. Neither the overall Diagnosis × Lists × Mediation interaction or the separate component $F$s approached accepted significance levels.

In general all groups showed some associative facilitation, improving their scores from List I to List II. The analysis of overall differences between lists yielded an $F$ of 11.80 ($p < .01$). Although the main Group × Lists interaction did not approach significance, the psychotic vs. nonpsychotic component attained a $p$ value close to .10. The direction of this trend is revealed in Table 1 where it will be noted that the degree of improvement from List I to List II decreases progressively from a difference in means of 3.84 for the psychiatrically normal to less than a trial (.87) for the schizophrenic patients.

Differences between psychotic and nonpsychotic patients in associative facilitation are clearly revealed when the course of learning is examined. Figure 1 shows the progress of learning for the three diagnostic groups on both List I and List II. Inspection reveals no differences between groups for List I. This was confirmed by an analysis of variance in which the Diagnosis × Successive Pairs interaction $F$ was less than 1.0. However, for List II the schizophrenic subjects displayed a tendency to take an increasing number of trials than either anxious or normal patients to learn successive pairs. The graphic picture was confirmed by an analysis of variance: the Diagnosis × Pairs $F = 2.33$, $p < .01$.

The analysis of the number of correct responses in the first five trials of List II was not undertaken directly, because of the potentially confounding effect of List I learning. It was decided to remove this variable through covariance analysis. However, the measure of List I learning to be used presented a problem: the correlation between the first five trials of List I and List II was low and not significant. Fortunately, mean trials to reach criterion on List I and the relevant

[2] The authors wish to thank A. W. Bendig who designed this analysis and Maryl Conwisher who assisted with the computation.

TABLE 3

MEAN NUMBER OF CORRECT RESPONSES IN THE FIRST FIVE TRIALS OF LIST II

| Diagnosis | Mediation | | Neutral | | Total | |
|---|---|---|---|---|---|---|
| | Actual[a] | Adjusted[b] | Actual[a] | Adjusted[b] | Actual[a] | Adjusted[b] |
| Normal | 11.27 | 11.57 | 6.87 | 7.00 | 18.13 | 18.56 |
| Anxiety | 9.27 | 8.61 | 5.80 | 5.70 | 15.07 | 14.43 |
| Schizophrenia | 8.87 | 9.17 | 4.27 | 4.23 | 13.13 | 13.13 |

[a] Refers to the actual mean.
[b] Indicates List II mean scores adjusted for the regression of the respective List I mediation, neutral, and total criterion scores.

Y variable yielded an $r$ of $-.60$. Requirements for homogeneity of regression were met with this measure, and it became the X variable in the subsequent analyses.

Three separate analyses of covariance were performed on the following scores: total correct responses, correct mediated responses, and correct neutral responses. The actual and adjusted means for each variable are presented in Table 3. The results of the analyses, presented in Table 4, reveal differences between patient populations in total correct responses. A test of normal and schizophrenic means yielded a $t$ of 2.48, $p$ of .02. A similar test of differences between anxious and nonpsychiatric patients resulted in a $t$ of 1.95, $p$ of .06.

The $F$ of 3.17 for correct neutral terms also warrants a test of individual means. In this case the differences between normal and schizophrenic subjects yielded a $t$ of 2.14, $p < .05$. The difference between the nonpsychiatric and anxiety group means was not significant.

TABLE 4

ANALYSIS OF COVARIANCE OF THE FIRST FIVE TRIALS OF LIST II

| Source | *df* | *MS* | *F* |
|---|---|---|---|
| Mediated pairs | | | |
| Diagnosis | 2 | 32.73 | 2.11 |
| Error | 41 | 15.53 | |
| Neutral pairs | | | |
| Diagnosis | 2 | 28.81 | 3.17* |
| Error | 41 | 9.08 | |
| Total list | | | |
| Diagnosis | 2 | 113.03 | 3.39** |
| Error | 41 | 33.30 | |

* $p < .06$.
** $p < .05$.

## DISCUSSION

The results of this experiment clearly demonstrate the facilitating effect of verbal mediation on learning in nonstudent populations. They also suggest that the mediation process is not fundamentally impaired in schizophrenia, and that theories resting on this assumption may be in error.

The findings relevant to the hypotheses of Mednick (1958) and Spence (1958) are complex, and require careful review. First of all, as Spence predicts, differences between groups appeared in the analysis of the "first few trials" of List II, rather than in the evaluation of mean trials to criterion.[3] It would appear that, despite the fact of overall associative facilitation, old response terms interfered differentially in the acquisition by the separate subgroups of the new pairs. The effect of these intruders was most marked for the schizophrenic subjects, as Mednick proposes, although the anxious subjects also showed impaired performance relative to normals. The prediction of an interference ef-

[3] Nevertheless, inspection of the individual criterion scores for the three groups was suggestive. Only one member of the normal group failed to achieve criterion on the mediated pairs of List II in fewer trials than on List I. Both the anxious subjects and the schizophrenic patients behaved comparably on mediated pairs: only two subjects from each of these groups gave similar evidence of negative transfer. However, the neutral pairs yielded quite different results: three normal, six anxiety neurotics, and nine schizophrenic patients (subgroups $n = 15$) learned List II in more trials than were needed to learn List I. This finding parallels the subgroup differences noted on the first few trials of List II. It also implies, at least in the case of neutral pairs, that List I learning interfered with List II learning, and furthermore that the interference effect was positively related to an increase in behavior pathology.

fect for high drive subjects on neutral pairs was realized for the schizophrenic group but not for the anxious subjects.

The fact that no difference between groups was apparent for mediated pairs is comprehensible in Spence's (1958) terms, if we consider the fact that two opposing factors are held to be operating. Although in this case the associative connections between the stimulus and its paired word were initially greater than with the other response terms (suggesting an advantage for anxious and schizophrenic subjects); subjects were also being asked to learn a response initially weaker than a potential intruder (i.e., the response learned 15 minutes earlier). It is reasonable to assume that advantage and disadvantage for high drive subjects cancel each other out, resulting in no differences between groups.[4]

The difference in learning curves between psychotic and nonpsychotic subjects is not readily integrated with the other findings. Spence's (1958) drive hypothesis suggests that differences should appear early rather than late in learning. The picture of increasing deviation with successive pairs implies the effects of a second variable.

The possibility that curve differences are an artifact of the experimental design was explored. The response terms for List II varied from subject to subject, and there might have been systematic differences between groups. As would be anticipated (Moran, 1953) the schizophrenic subjects tended to report less common associations than did the normal or anxious subjects. It is conceivable that these were more difficult to learn, and this affected the rate at which the subject reached criterion on successive pairs. In point of fact, this variable might have contributed to subgroup differences on the other measures already discussed. However, a Pearson $r$ (sample $N = 18$) between the number of pairs correct in List II and an index of deviation based on the Thorndike-Lorge word count was insignificant and in the direction which contradicts this hypothesis.

Two possibilities remain which could account for the findings. The schizophrenics might have suffered from unique fatigue effects or flagging motivation. However, this hypothesis seems difficult to accept in the light of their failure to show this phenomenon on List I, and the fact that List II learning was slightly more rapid, and followed a 15-minute rest period. A second, more likely possibility was suggested by the behavior of the psychotic group during learning. Late in learning, these subjects frequently responded to a stimulus syllable with a response term which had already achieved criterion and been dropped from the list. This happened rarely with the other groups. For these latter subjects, when a response was eliminated its status as a potential intruder was markedly reduced. The schizophrenics, on the other hand, behaved as if nearly all the competitors remained active, if not phenomenally present, until the very end of the list. In List I the effects might be minimal. However, in List II, where a neutral stimulus was always more closely related to some response other than the one with which it was paired, and intruders from List I could reassert themselves, this might have greatly retarded the learning of the final few pairs.

The above interpretation raises the question of the extent to which subjects were "conscious" of mediators or interfering terms during the learning. Previous investigators (Bugelski & Scharlock, 1952; Foley & Cofer, 1943) have expressed considerable interest in this issue; however, an adequate method of determining awareness has not been developed.

In the present research, the last 5 subjects seen in each diagnostic group were closely questioned at the end of the experimental session, concerning their thoughts and impressions during learning. Only 2 of the 15 subjects interviewed suggested that they knew the List II response terms to be their own associations. Eight of the 15 said that they were aware of a relationship between the two lists, although the connection was usually

[4] It is also possible that differences failed to appear because the psychiatrically disturbed patients offered unique associations which were systematically different in their ability to mediate generalization. It might be hypothesized that these associations were less stable, less able to mediate generalization, and the effect of drive variables was simply to raise the learning of anxious subjects to the same level of proficiency as that displayed by the normal patients.

poorly understood. All 7 who were totally unaware of facilitating or interfering associations displayed the positive effects of the mediated connections in List II learning. This latter finding is consistent with previous research, which suggests that awareness is not necessary to verbal mediation.

The distribution of awareness by diagnostic groups is of some interest. Four schizophrenics, two anxious, and one normal subject were unaware of connecting associations. Thus, insight seems to follow the pattern of interference effects. It is possible that while awareness does little to facilitate mediated association, it may be an important factor in keeping out intruders. Subjects who are aware that a compelling connection is nevertheless wrong, are conceivably in a better position to ignore that association than a subject who fails to perceive any relationship at all.

In discussing the performance of schizophrenic subjects on a psychomotor task, the senior author suggested that deficit in performance was attributable to a failure of "set" or "self-instruction"—an inability to discriminate relevant cues during performance and reprogram the organism to take account of them (Cavanaugh, Cohen, & Lang, 1960; Lang, 1959). The lack of awareness displayed by the schizophrenics on the mediation task suggests a similar process, which may have contributed both to retardation in the course of learning and the deficit noted on neutral and total pairs in the first five trials of List II.

In the absence of a sampling design attesting to the representativeness of our samples of schizophrenic and other populations, our results bear primarily on the mediation process in association rather than on differences between types of subjects. However, the present findings strongly suggest that special transfer and interference phenomena are important contributors to the symptom picture in chronic schizophrenia.

## SUMMARY

Schizophrenics, anxiety neurotics, and psychiatrically normal patients ($N = 45$) learned two lists of 10, syllable-word paired associates. Stimulus terms were identical on Lists I and II. Half of the List II pairs had response terms which were associates of their List I responses; the response terms of the other five (neutral) pairs were not related to their stimuli by any such mediating association.

The entire population showed improvement from List I to List II, and greater improvement for mediated than for neutral pairs. Schizophrenic subjects took progressively more trials than the other two groups to learn successive pairs in List II. Both the anxiety neurotics and the schizophrenics had fewer correct responses from the total list in the first few trials of List II than did normal subjects. Schizophrenics also responded with fewer correct neutral pairs than the normals.

## REFERENCES

BENDIG, A. W. The development of a short form of the manifest anxiety scale. *J. consult. Psychol.*, 1956, **20**, 384.

BLEULER, E. *Dementia praecox or the group of schizophrenics*. New York: International Univer. Press, 1950.

BUGELSKI, B. R. *The psychology of learning*. New York: Holt, 1956.

BUGELSKI, B. R., & SCHARLOCK, D. P. An experimental demonstration of unconscious mediated association. *J. exp. Psychol.*, 1952, **44**, 334–338.

CAVANAUGH, D. K., COHEN, W., & LANG, P. J. The effect of "social censure" and "social approval" on the psychomotor performance of schizophrenics. *J. abnorm. soc. Psychol.*, 1960, **60**, 213–218.

DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy: An analysis in terms of learning, thinking and culture*. New York: McGraw-Hill, 1950.

FOLEY, J. P., & COFER, C. N. Mediated generalization and the interpretation of behavior: II. Experimental study of certain homophone and synonym gradients. *J. exp. Psychol.*, 1943, **32**, 168–175.

GOLDSTEIN, K. The significance of special mental tests for diagnosis and prognosis in schizophrenia. *Amer. J. Psychiat.*, 1939, **96**, 575–588.

GOSS, A. E. Mediating responses and concept formation. *Amer. Psychologist*, 1958, **13**, 391. (Abstract)

GRIFFITH, B. C., SPITZ, H. H., & LIPMAN, R. S. Verbal mediation and concept formation in retarded and normal subjects. *J. exp. Psychol.*, 1959, **58**, 247–251.

HANFMANN, EUGENIA, & KASANIN, J. Conceptual thinking in schizophrenia. *Nerv. ment. dis. Monogr.*, 1942, No. 67.

KITAO, N. Effects of mediated association on paired-associate learning: I. Facilitation effect and interference effect. *Jap. J. Psychol.*, 1960, **31**, 1–6.

LANG, P. J. The effect of aversive stimuli on reaction time in schizophrenia. *J. abnorm. soc. Psychol.*, 1959, **59**, 263–268.

MEADOW, A., GREENBLATT, M., & SOLOMON, H. C. "Looseness of association" and impairment in abstraction in schizophrenia. *J. nerv. ment. Dis.*, 1953, **118**, 27–35.

MEDNICK, S. A. A learning theory approach to research in schizophrenia. *Psychol. Bull.*, 1958, **55**, 316–327.

MORAN, L. J. Vocabulary knowledge and usage among normal and schizophrenic subjects. *Psychol. Monogr.*, 1953, **67**(20, Whole No. 370).

OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.

RUSSELL, W. A., & STORMS, L. H. Implicit verbal chaining in paired-associate learning. *J. exp. Psychol.*, 1955, **49**, 287–293.

SPENCE, K. W. A theory of emotionally based drive (*D*) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.

(Received January 23, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 121–129

# A DISSONANCE THEORY APPROACH TO DEFENSIVE PROJECTION [1]

DANA BRAMEL

*University of Minnesota*

Ego defensive processes, as discussed in psychoanalytic theory, often seem to bear some resemblance to the cognitive changes dealt with in Festinger's (1957) dissonance theory. This observation has led to a comparison of the two theories and to consideration of the possibility that certain of the Freudian defense mechanisms might occur in response to dissonance. Especially interesting from the point of view of social psychology is the concept of projection, since it clearly has implications for interpersonal relations.

Consider those situations, described in psychoanalytic theory, in which the individual's perception of some aspect of himself is contrary to his internalized standards of right and wrong (the superego). According to the theory the perception of this information arouses fear of punishment, perhaps especially a fear of painful guilt feelings (Fenichel, 1945). In order to avoid further anxiety and guilt feelings, the ego is said to initiate defensive measures.

In order to determine the relevance of dissonance theory to this phenomenon, one must ask whether dissonant relations would be expected to exist among the cognitions involved. Imagine, for example, a person who considers homosexuality a bad and disgusting thing; on some occasion he is suddenly exposed to information strongly implying that he has homosexual tendencies. According to classical psychoanalytic theory, the crucial relation is the conflict between the information or impulse and the demands of the superego. Is the new cognition—that one has homosexual tendencies—necessarily dissonant with one's belief that such tendencies are bad and that one should not have them? The answer is no. People who conceive of themselves as possessing a mixture of good and bad traits, or a preponderance of bad traits, would not generally expect that new information would be favorable to themselves or consistent with superego standards. It follows that a discrepancy between the new information and superego standards, although threatening in the psychoanalytic sense, would not necessarily be dissonant.

This is not to say, however, that dissonance would be completely absent from the cognition of the person who recognizes that he is not perfect. For example, when he discovers he has homosexual tendencies, this knowledge may be dissonant with his specific belief that he is really quite masculine, even though it is not dissonant with his conviction that homosexuality is a bad thing.

The point can be clarified by a hypothetical example. Imagine two people, A and B. Both consider homosexuality a very bad thing, and both believe they are quite lacking in such motivation. A believes he is an extremely fine person in general; B sees himself as possessing almost no favorable characteristics. Both are then confronted with information that they have strong homosexual tendencies. For both this information is contrary to superego standards and dissonant with their belief that they are not homosexual. However, for A the information is also strongly dissonant with his belief that he is a nearly perfect person. For B, on the other hand, the information is quite consonant with his belief that he is a failure. The new information produces more dissonance for the person with high self-esteem, even though the conflict with the superego is substantially the same for the two people. This emphasis upon the actual self-concept in the dissonance theory ap-

[1] This article is based upon a doctoral dissertation submitted to Stanford University. The author wishes to express his indebtedness to Leon Festinger for valuable assistance throughout the research. Advice and criticism from Robert R. Sears, Nathan Maccoby, Quinn McNemar, and Stanley Schachter are gratefully acknowledged. Special thanks are also due Alvin I. Haimson for assistance in conducting the experiment.

proach reveals a difference in focus of the two theories.

Can dissonance involving the self-concept be reduced by projecting the offending trait onto other people? Perhaps the most effective mode of dissonance reduction would be to deny the implications of the information. Let us assume, however, that the information is so unambiguous that successful denial is not possible for the person. He is compelled to ascribe the undesirable trait to himself. Under these circumstances, attributing the trait to other persons might reduce dissonance in several ways. By attributing it to respected people, the projector may enable himself to re-evaluate the trait. If respected persons possess it, then perhaps it is not so bad a thing after all. Then possession of the trait would not be contradictory to a favorable level of self-esteem. Another possibility is that the person may attribute the trait to members of his reference or comparison group (Festinger, 1954). In this way he could convince himself that he does not deviate from the persons with whom he ordinarily compares himself. If he is only average in his possession of the trait, then subjectively his favorable level of self-esteem is not so strongly negated.

These possibilities suggest that indeed projection may be used as a means of reducing dissonance. There are several studies in the literature (for example, Backman & Secord, 1959 unpublished; Murstein, 1956; Wright, 1942) which are specifically relevant and show positive results, but all leave certain important issues unresolved. Consequently, the experiment reported here was conducted to test whether projection occurs in response to dissonance and to throw some light on the particular ways in which this attribution may reduce dissonance.

The hypotheses to be tested were these:

1. If a person is exposed to information strongly implying that he possesses an undesirable characteristic, he is more likely to attribute the trait to others if the information is dissonant with his level of self-esteem; the greater the dissonance, the more likely it is that projection will occur.

2. If a person is compelled to ascribe an undesirable and dissonant characteristic to himself, he will be motivated to attribute the characteristic to favorably evaluated persons and/or to persons with whom he ordinarily compares himself.

## METHOD

*Overview.* Subjects in the Favorable condition received falsified personality test results aimed at temporarily increasing their general level of self-esteem; subjects in the Unfavorable condition received parallel information intended to lower their general self-esteem. Subsequently, all subjects were privately exposed to further falsified information of an undesirable nature about themselves. It was hypothesized that this information, being more dissonant with the self-concept of subjects in the Favorable condition, would lead to more projection in that condition. Attribution was measured by asking each subject to rate another subject with whom he was paired.

*First session.* Each subject who signed up for the experiment appeared individually for the first session. He was told that the first part of the experiment was designed to discover what kinds of people had insight into themselves. He was asked to take a number of personality tests, which, he was told, would be carefully and confidentially analyzed by three members of the clinical psychology staff. He was informed that, after the tests were scored, he would learn the results in an interview, during which time his self-insight was to be measured. Among the tests included were the Taylor Manifest Anxiety scale, the *F*, *K*, and *L* scales from the MMPI, and an adjective checklist self-concept measure.

At the end of the hour, the subject was told that the second session would also include a measure of his ability to judge the personality of another person on the basis of a first impression.

*Second session.* On the basis of the self-concept measure subjects were paired for the second session by matching their level of self-esteem and their concept of their own masculinity.

At the beginning of the second session the two unacquainted subjects scheduled for the hour were introduced to each other. In order to aid subjects in forming an impression of each other, the experimenter asked each in turn (in the presence of the other) a set of questions about himself and his attitudes toward certain current events. At the conclusion of this meeting the subjects were separated and interviewed privately regarding the results of their personality tests.

Unknown to the subject, the "results" which he received had been prepared with no reference whatsoever to his actual test performance. There were only two test reports used in the experiment, one very favorable, and the other very unfavorable. The reports covered the personality "dimensions" of creativity, hostility, egocentricity, and over-all maturity. Each section of the report gave a rather detailed discussion of the test results bearing upon

the particular dimension. The tone was objective and the general favorability was very consistent throughout the report. The two reports were very similar in form, but the specific contents were directly opposite in implication.

After having been assigned randomly to his experimental condition, the subject was read the report in private by the experimenter, and its discrepancies from the subject's present self-concept were explicitly pointed out by the experimenter. The report was finished in approximately 20 minutes. In each pair of subjects, one was assigned to the Favorable condition, one to the Unfavorable condition. Two interviewers were used for this part of the experiment, alternating between the two conditions.

Following the test report, the two subjects were brought together in another room, where they expected to make some personality judgments about each other. Each was then given a questionnaire consisting of 11 polar adjective seven-point scales, to be used to rate the other person. An over-all favorability score could be computed across the scales. Examples were masculine-feminine, friendly-hostile, competent-incompetent, and mature-immature. A self-concept measure followed, consisting of 16 polar adjective scales similar to those included in the prior rating of the other person. These were selected partially to tap dimensions covered in the personality reports, and to serve as a check on the manipulation of self-esteem. As in the previous set, they included the item, masculine-feminine, and could be summated to provide a general favorability score. Emphasis was placed upon the anonymity of the questionnaires and upon the earnest request that the subject respond "as *you* see yourself, from your own point of view at the present time."

*Introducing the undesirable cognition.* It was expected that informing a male undergraduate that he has homosexual tendencies would be sufficiently dissonant under certain conditions to provoke defensive behavior. Care was taken to ensure that the degree of threat was not extreme and that no damaging effects would remain at the end of the experiment. These precautions will be discussed in more detail below.

At this point in the experiment, while making ratings of partner and self, the two subjects were seated along one side of a long table, separated by about 4 feet, both facing a projection screen 6 feet in front of them. On the table in front of each subject was a small plywood box containing a galvanometer dial facing him. Issuing from the box were two wires with electrodes on the ends. Each box, with its dial, was completely shielded from the other persons in the room. Thus, each subject perceived his apparatus immediately in front of him and could not see the other subject's apparatus.

Next, the experimenter read a set of instructions to set the stage for the undesirable cognitions about homosexuality. These instructions were largely of a deceptive nature. It was explained that this part of the experiment would be concerned with the perception of sexual arousal. An elaborate explanation of the physiology of sexual arousal and the sensitive techniques for its measurement followed. Care was taken to distinguish the galvanometer response to sexual arousal from that commonly associated with anxiety reactions, by pointing out the unmistakable signs of the former. Considerable emphasis was placed on the unconscious nature of sexual arousal and the impossibility of exerting conscious control over its expression in the "psychogalvanic skin response." It was further explained that the experimenter was investigating the perception of homosexual rather than heterosexual arousal. The task set for the subject was to observe his own sexual arousal response on his galvanometer for each of a series of photographs of men which would be projected onto the screen. He was to record this figure on a page of a small anonymous booklet. After he had recorded his own arousal level for the particular picture on the screen, he was to make an estimate of the needle indication of the other subject's apparatus for the same photograph. All subjects were explicitly told that movements of the dial would indicate homosexual arousal to the photographs. As a precaution against excessive threat, they were told that persons with very strong homosexual tendencies would consistently "go off the scale." Further, the anonymity and privacy of the situation were carefully spelled out, with the intention of convincing the subject that no one but he would know what his own responses had been.

Unknown to the subject, the supposed "psychogalvanic skin response apparatus" was not actually responding to changes in his own level of sexual arousal to the pictures. Rather, the galvanometers in each of the two boxes were controlled remotely by the experimenter. Concealed wires led from the galvanometers, in a direct current series circuit, to a calibrated variable resister. Thus, the experimenter exerted complete control over the movements of the needles, which were identical for the two subjects. Each photograph had been assigned an "appropriate" needle reading in advance, so that those depicting handsome men in states of undress received more current than did those depicting unattractive and fully clothed men. Both subjects were, thus, led to believe that they were sexually aroused by certain pictures and not by others, according to a consistent pattern. Both subjects were confronted with exactly the same stimulus input at this point of the experiment.[2]

It was expected that the instructions would be so impressive to the subject that denial of the fact that homosexual arousal was being indicated would be very difficult. By closing off certain alternative avenues of dissonance reduction, such as the cognition that the apparatus was untrustworthy, it was intended that the situation would be conducive to the appearance of defensive projection. According to the hypothesis, subjects in the Favorable con-

[2] A similar experimental technique was independently devised by Harold Gerard (cf. Gerard, 1959; Gerard & Rabbie, 1960).

dition should experience considerable dissonance when observing their needle jump in response to photographs of attractive males. For subjects in the Unfavorable condition there would be many cognitions consonant with the new information concerning homosexuality, and not so many dissonant cognitions. For most subjects there would, no doubt, be some dissonance due to their prior belief that they are not homosexual, but the two conditions would not differ in this respect. Since the test reports contained no material concerning sexuality, subjects in the two conditions were not expected to differ in their concepts of their own masculinity or in their superego standards.

Fifteen photographs of men were used. Many of the men were almost entirely nude and had physiques somewhat more delicate and posed than those typically found in physical culture magazines. These photographs were chosen on the assumption that subjects might perceive them as being the type toward which homosexuals would be attracted.

*Measuring attribution.* It seemed that the most meaningful measure for testing the hypotheses would be a score representing the difference between the subject's own recorded score and his estimate of his partner's galvanic skin response. This should most accurately reflect the subject's comparison between himself and his partner. Therefore, a total score was computed for each subject, taking the algebraic sum of the differences between own and attributed scores across the 15 photographs. This summary score (*P* score) would be positive if the subject attributed (on the average) higher needle indications to his partner than to himself (i.e., attributed greater homosexual arousal). It would be zero if on the average there was no difference between own and attributed scores. It would be negative if the subject attributed lower needle indications to the other subject than to himself.

Following the threatening material, the subject responded to anonymous questions about his own and his partner's degree of possession of homosexual tendencies, and about his attitude toward the "psychogalvanic skin response" as a measure of such tendencies.

A considerable amount of time at the end of the experiment was allocated to explaining the true nature of the study and demonstrating in detail that the personality reports and the apparatus were incapable of giving a correct evaluation of a person. The expression of relief which often followed the unveiling of the deceptions indicated that the manipulations had been effective. The necessity for the deceptions used in the experimental analysis of such delicate processes was carefully explained, and all questions were answered. Not until the subjects seemed quite restored and satisfied was the session ended. All available evidence indicates that the subjects considered the experiment interesting and worthy of their participation.

*Subjects.* All subjects in the experiment were undergraduate men registered in the introductory psychology course. Not all of those who took part in the first session were selected to finish the experiment. Those who scored very high on the Taylor Manifest Anxiety scale and at the same time very low on the defensiveness scale of the MMPI were excluded from the second session, since there was the possibility that the manipulations might be too threatening for them. Of the 98 subjects who participated in both sessions of the experiment, 14 were excluded from the analysis—8 for suspicion regarding the procedure, 3 for excessive age (over 30), and 3 for failure to obey the instructions. Of those excluded, 7 were in the Favorable and 7 in the Unfavorable condition.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF VARIABLES MEASURED PRIOR TO INTRODUCTION OF THE UNDESIRABLE COGNITION

| Variable | Favorable ($n = 42$) | Unfavorable ($n = 42$) |
|---|---|---|
| Before self-esteem manipulation | | |
| Initial self-esteem (checklist) | | |
| *M* | 14.5 | 15.1 |
| *SD* | 3.6 | 3.3 |
| After self-esteem manipulation | | |
| Self-esteem (seven-point scales) | | |
| *M* | 5.52 | 4.20 |
| *SD* | .54 | .86 |
| Favorability of rating of partner | | |
| *M* | 4.79 | 4.90 |
| *SD* | .69 | .68 |
| Rating of masculinity of partner | | |
| *M* | 5.39 | 5.30 |
| *SD* | 1.11 | 1.32 |
| Rating of own masculinity | | |
| *M* | 5.87 | 5.56 |
| *SD* | 1.00 | 1.02 |

## RESULTS

*Adequacy of experimental operations.* The major independent variable was the level of self-esteem, or the number of favorable self-referent cognitions. A checklist measure of self-esteem administered before the manipulations revealed no initial difference between the groups. The effectiveness of the personality reports was determined by comparing the Favorable and Unfavorable groups on self-esteem as measured by adjective scales shortly after the manipulation. Mean favorability scores could range from a low of 1 to a high of 7. The results are shown in Table 1. The mean for the Favorable group was significantly higher than for the Unfavorable group ($t = 8.35$, $p < .001$).[3] We may infer, therefore, that the desired difference in self-esteem was successfully created by the fraudulent test reports.

Another important problem of experimental

[3] All reported significance levels are based upon two-tailed tests.

control had to do with the favorability of the subjects' ratings of each other prior to introduction of the cognitions concerning homosexuality. A score was calculated for each subject, taking the mean of his ratings of his partner (scored for favorability) across 10 polar adjective scales, excluding the item, masculine-feminine (considered separately below). The first half vs. second half reliability of the score was .57. Possible scores could range from 1 (very unfavorable) to 7 (very favorable). As shown in Table 1, there was no significant difference between the two conditions in favorability of rating of partner.

On the masculine-feminine scale, 1 indicated "very feminine" and 7 indicated "very masculine." Mean scores of the two groups did not differ on this scale, as shown in Table 1. Further, the groups did not differ significantly in their rating of their own "masculinity-femininity."

On the basis of these comparisons, it seems justifiable to conclude that the Favorable and Unfavorable groups did not differ regarding these possible artifactual effects of the self-esteem manipulations.

*Self-esteem and projection.* Before comparing the two experimental conditions on the attribution of homosexual arousal, let us check the reliability of the measure, the *P* score. For the first 23 subjects used in the experiment, the series of 15 photographs was repeated, yielding a set of 30 judgment situations for each subject. The discrepancies between his own recorded dial readings and his estimates of his partner's dial were summed separately for the first 15 and the second 15 exposures of the photographs. The correlation between these two sums (*P* scores) was .95. For subsequent analyses, only the *P* score for the first 15 photographs was used.

Subjects in the two conditions did not differ, on the average, in their own recorded scores. That is, they were equal in the accuracy with which they recorded their own needle indications. Therefore, the *P* scores, which were partially derived from the subjects' own recorded dial readings, could not differ between the two groups simply as a function of differences in own recorded scores.

For evidence concerning the relation between dissonance and projection, let us look

TABLE 2

MEANS AND STANDARD DEVIATIONS OF ATTRIBUTION MEASURED AFTER MANIPULATION OF SELF-ESTEEM

| Attribution | Favorable ($n = 42$) | Unfavorable ($n = 42$) |
|---|---|---|
| Raw *P* score | | |
| *M* | −2.95 | −11.45 |
| *SD* | 25.52 | 27.96 |
| *P* score adjusted for prethreat judgment of masculinity of partner | | |
| *M* | +4.65 | −4.76 |
| *SD* | 24.28 | 23.99 |

first at the gross differences between the experimental conditions. As shown in Table 2, the mean *P* score for the Favorable condition was −2.95; the mean for the Unfavorable condition was −11.45. Thus, subjects in the high dissonance condition tended to say that their partner's arousal level was about the same as their own, while those in the low dissonance condition tended to say their partner's arousal was somewhat less than their own. Attributing one's own characteristic to others was therefore more frequent in the high dissonance, or Favorable, group. The difference between the means yielded a $t$ (for correlated means, due to matching) of 1.52, which is at the .13 level of significance. In order to arrive at a more firm conclusion regarding the outcome of this comparison, let us look at another source of variance which can be taken out of the gross variance in the *P* scores.

It had been anticipated that part of the variance in *P* scores would be due to the impression of masculinity created by the partner prior to the introduction of the undesirable cognition. If the subject rated his partner as very masculine on the masculine-feminine scale, then he would be likely to make somewhat lower (less homosexual) needle estimates for his partner than if he had rated him as very feminine. Correlations were therefore calculated between perceived masculinity of the partner (prethreat) and attribution of homosexual arousal to the photographs (*P* score). Within the Favorable group the resulting product-moment correlation was −.32 ($p < .05$); in the Unfavorable group the correlation was −.52 ($p < .01$). It will be remembered that the two groups did not differ in their mean (prethreat) rating of the

masculinity of the partner (as measured by the masculine-feminine scale); in addition, the distributions of these ratings were very similar in the two groups. Therefore, it was decided that the rather similiar within-groups correlations would justify combining the groups, calculating the correlation between the two variables in the total sample, and computing adjusted $P$ scores as deviations from the regression line. By means of this procedure the variance associated with how masculine the partner appeared (*prior* to the threat) could be partialed out. Within the total sample ($N = 84$) the correlation between "masculinity" (prethreat) and $P$ score was $-.42$. Each subject's $P$ score deviation from the regression line was calculated, and the resulting scores were then interpreted as reflecting differences in attribution due to factors other than the initial perceived masculinity of the partner. The adjusted means for the groups are shown in Table 2. A $t$ test for correlated means yielded a $t$ of 2.04 ($df = 40$), significant beyond the .05 level. On the basis of these results one may conclude that the groups differed in attribution in the direction predicted by the hypothesis relating dissonance and projection.

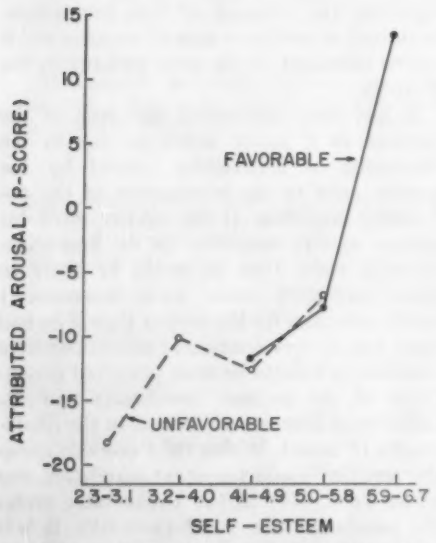



FIG. 1. Mean attribution of homosexuality as a function of level of self-esteem.

In all subsequent comparisons, the original, unadjusted $P$ scores were used. The use of the simpler score should make interpretations clearer and more direct, especially in the case of within-conditions analyses.

It is of interest to look at the relation between self-esteem and attribution of arousal within the two experimental conditions. The product-moment correlation within the Favorable group was $+.29$ ($p < .07$); within the Unfavorable group the correlation was $+.10$ ($p < .55$). The relations are presented graphically in Figure 1, showing the mean raw $P$ scores within each condition as a function of increasing self-esteem. The lack of correlation in the Unfavorable condition suggests that projection as a means of reducing dissonance occurred only when the amount of dissonance was quite high. From the point of view of the theory, this is not surprising. For the person with low self-esteem the undesirable information is actually consonant with his general self-evaluation (although dissonant with his specific cognitions about his adequate masculinity). Only for the person who believes he is consistently good will the undesirable information be strongly dissonant with his general self-esteem.

In Figure 1 it can be seen that the two groups show considerable continuity where they overlap in level of measured self-esteem. This fact is important because it implies that the self-esteem manipulations did not have strong opposed artifactual effects upon the amount of attribution of homosexual arousal. If the personality test report interviews had had effects on attribution in ways other than through the self-esteem variable, then differences between conditions might have appeared when considering subjects in the two groups with equivalent levels of self-esteem. Judging from Figure 1, persons in the two groups who had equivalent measured self-esteem levels apparently reacted similarly to the undesirable cognition.

*Projection and attitude toward available social objects.* There is a well known judgmental tendency which leads a person to perceive others as possessing traits consistent with his general evaluation of those others (a halo effect). On the basis of the halo effect alone, one would expect a tendency to at-

tribute homosexuality (an unfavorable trait) to persons who are evaluated in general relatively negatively. However, the presence of dissonance resulting from self-ascription of homosexuality should introduce a contrary tendency. To the extent that projection, of the type defined in this report, occurs, it should be aimed primarily at persons who are relatively favorably evaluated. Since in this experiment projection was expected to occur to a greater extent in the Favorable group, it follows that the empirical pattern of attribution in that condition should be some compromise between the projection pattern and the halo pattern, since one may expect both forces to be operating. In the Unfavorable condition, on the other hand, one would expect to find a pattern more closely resembling the pure halo pattern, due to the absence of large amounts of dissonance.

In Figure 2 the results are shown separately for the two conditions. Mean *P* scores are shown as a function of increasingly favorable evaluation of the partner, as measured independently and prior to the introduction of the cognitions regarding homosexuality. It can be seen that the results are consistent with the hypothesis. For relatively negative and moderate levels of evaluation of the partner, subjects in the Favorable and Unfavorable conditions attributed homosexual arousal consistent with a halo effect. The less favorably they rated the other subject, the more homosexuality they attributed to him. However, when the partner had been evaluated very favorably, subjects in the two experimental conditions reacted in quite different ways. The Unfavorable group continued to follow the halo pattern, attributing very low homosexuality to the partner. The Favorable group, in contrast, exhibited no decrease in attribution when confronted with favorably evaluated objects. In fact, there was a slight but insignificant increase. The difference between the mean *P* scores of the Favorable ($n = 9$) and Unfavorable ($n = 12$) groups at the high respect point (5.4 to 6.4 in Figure 2) was significant beyond the .05 level by the $t$ test. Since the subject's respect for his partner was not experimentally manipulated, it is possible that via self-selection other variables may be contributing to the observed difference. It should be pointed out, for example, that the measured self-esteem level of subjects in the Favorable group (with respected partner) was slightly higher than that of other subjects in the Favorable group, so that some of the tendency of these particular subjects to project may be traceable to their higher self-esteem rather than to their attitude toward their partner. Unfortunately, the number of cases is too small to allow an internal analysis to throw light on this question. All things considered, one can be fairly confident that the difference between the conditions does reflect a tendency for projection to be directed toward favorably evaluated persons under these circumstances.

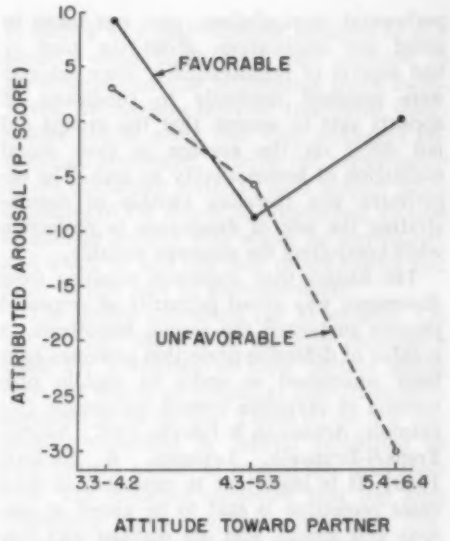



FIG. 2. Mean attribution of homosexuality as a function of favorability of attitude toward partner.

## DISCUSSION

The results provided good support for the central hypothesis, that projection can be a response to dissonance involving the self-concept. Subjects in the Favorable condition, for whom the undesirable information about homosexuality was more dissonant with the self-concept, attributed more arousal to other persons. It is very unlikely that this difference was due to differences between conditions in severity of superego standards concerning homosexuality. In designing the ex-

perimental manipulations, care was taken to avoid any implications about the good or bad aspects of homosexuality. Since subjects were assigned randomly to conditions, it appears safe to assume that the groups did not differ on the average in their moral evaluation of homosexuality as such. The experiment was therefore capable of demonstrating the role of dissonance in projection while controlling the superego variable.

The finding that projection resulting from dissonance was aimed primarily at respected persons supported the second hypothesis. A number of defensive projection processes have been summoned in order to explain phenomena of prejudice toward out-groups (for example, Ackerman & Jahoda, 1950; Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950). It is important to note that in most cases projection is said to be aimed at persons and groups who are disliked and considered incomparable and inferior to the projector. This is, of course, quite different from the kind of projection revealed in the present experiment.

Let us consider possible alternative explanations for the results shown in Figure 2. Perhaps the dissonance introduced by the undesirable information in the Favorable condition led directly to a re-evaluation of homosexuality. That is, perhaps these subjects were able to change their attitude toward the trait without the intermediate step of associating it with favorably evaluated persons. This re-evaluation prior to attribution might then affect the pattern of attribution in such a way as to give the appearance, deceptively, of projection. Once the undesirability of the trait was reduced, there would be less tendency to attribute it differentially to disliked persons (halo effect). The effect of such a process would be to attenuate the halo pattern in the Favorable condition, and might be revealed in part as a greater tendency to attribute homosexuality to respected persons, as compared with subjects in the Unfavorable condition. Is this hypothesis capable of explaining the results of the experiment without resort to the hypothesis of defensive projection?

The data show that subjects in the Favorable condition did tend to follow the halo pattern when the partner was evaluated unfavorably or moderate in favorability. In Figure 2, the difference between the unfavorable ($n = 9$) and moderate ($n = 24$) points within the Favorable group was significant beyond the .10 level by the $t$ test, and was at least as striking as the pattern for the Unfavorable group. If in fact the high dissonance in the Favorable condition had led directly to re-evaluation of homosexuality (without projection), then the halo effect would have been attenuated in this condition at all points in Figure 2. There is no apparent reason for supposing that re-evaluation of the trait (without defensive projection) would have occurred among Favorable subjects confronted with respected partners more so than among other subjects in the Favorable group.

Consider another alternative explanation of the results. Perhaps the personality test report led subjects in the Favorable condition to conclude that they were generally better than other people, while those in the Unfavorable condition concluded they were worse than other people. Both groups of subjects then were given information that they possessed some degree of homosexual arousal. If subjects in the Favorable condition believed they were generally superior people, they might then have deduced that their partner was likely to be less worthy than themselves. This could result in rating the partner as possessing a greater amount of homosexual motivation. Persons in the Unfavorable condition would, by similar reasoning, conclude that their partner possessed a smaller amount of homosexual arousal. Such nondefensive processes could account for the over-all difference in attribution between the two experimental conditions.

The data, as presented in Figure 2, cast doubt upon this alternative explanation. If the over-all difference between conditions were due to nondefensive deductions from the personality test reports, one would expect to find differences between the Favorable and Unfavorable conditions at all points along the attitude-toward-partner dimension. It is apparent, on the contrary, that the difference between conditions occurred only when the partner had been rated favorably.

It is interesting to speculate about the conditions under which dissonance with the self-concept will lead to projection onto favorably evaluated persons. In this experiment the dissonance producing information was probably quite striking and unambiguous to the subjects. With great care the experimenter had explained that movement of the needle in response to looking at the photographs was a clear and indisputable sign of homosexual arousal. The situation was such that outright denial of the meaning of the needle movements would have been quite difficult for persons in reasonable touch with reality. It is very likely that these subjects were forced to accept the information as implying some degree of homosexual arousal in themselves. Under these circumstances of self-ascription, a good way to reduce the dissonance remaining was to try to get desirable people "into the same boat."

However, if the subjects had been able to deny the direct implications of the galvanic skin response, then a different pattern of attribution might have been observed. If the information were sufficiently ambiguous, so that partial denial occurred, then it would no longer be so comforting to attribute the undesirable trait to persons with whom the subject ordinarily classes himself. That is, when one is attempting to avoid self-ascription, it probably does not help to ascribe the trait to others who are seen as generally similar to one's self. Whether, when denial is possible, projection tends to be directed toward undesirable persons or out-groups, is an interesting question for further experimental exploration.

## SUMMARY

A laboratory experiment was conducted to investigate some of the conditions affecting the occurrence of defensive projection. It was hypothesized that such projection is a positive function of the amount of cognitive dissonance resulting from the introduction of a self-referent cognition of negative valence. Further, it was hypothesized that certain types of defensive projection are likely to be selectively aimed at persons who are favorably evaluated by the threatened individual. Two groups of normal subjects were prepared in such a way that different amounts of dissonance would result from their exposure to the same undesirable information about themselves. All subjects received fraudulent information to the effect that they possessed homosexual tendencies. During the presentation of the disturbing material, each subject made estimates of the degree of homosexual arousal of another subject with whom he was paired, and whom he had met only rather briefly just prior to this part of the experiment.

The results supported the hypotheses. On the average, subjects in the high dissonance condition attributed to their partner about the same degree of arousal as they themselves appeared to be having. Those in the low dissonance condition in general attributed to their partner a level of arousal less than their own. The evidence suggested that the high dissonance group projected only when confronted with a partner whom they had previously evaluated quite favorably on adjective rating scales.

The relation between the psychoanalytic and dissonance theory approaches to defensive processes was discussed. It was proposed that study of the selection of objects should throw light upon the possible existence of distinct varieties of defensive projection.

## REFERENCES

ACKERMAN, N. W., & JAHODA, MARIE. *Anti-Semitism and emotional disorder.* New York: Harper, 1950.

ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality.* New York: Harper, 1950.

FENICHEL, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.

FESTINGER, L. A theory of social comparison processes. *Hum. Relat.*, 1954, **7**, 117–140.

FESTINGER, L. *A theory of cognitive dissonance.* Evanston, Ill.: Row, Peterson, 1957.

GERARD, H. B. A specific study: Conflict and conformity. *Amer. Psychologist*, 1959, **14**, 413. (Abstract)

GERARD, H. B., & RABBIE, J. M. Fear, affiliation, and social comparison. *Amer. Psychologist*, 1960, **15**, 409. (Abstract)

MURSTEIN, B. I. The projection of hostility on the Rorschach and as a result of ego-threat. *J. proj. Tech.*, 1956, **20**, 418–428.

WRIGHT, BEATRICE A. Altruism in children and the perceived conduct of others. *J. abnorm. soc. Psychol.*, 1942, **37**, 218–233.

(Received January 25, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 130-135

# THE INFLUENCE OF HUNGER ON THE LEARNING AND RECALL OF FOOD RELATED WORDS [1]

SEYMOUR EPSTEIN AND HERBERT LEVITT

*University of Massachusetts* — *University Hospitals, University of Wisconsin*

Despite a number of studies on the steering or directive effects of primary drives [2] on perception, association, and imagination (see, for example, Clark, 1952; Epstein & Smith, 1957; McClelland & Atkinson, 1948; Murray, 1959; Wispé, 1954), not a single such study has been reported on learning in humans. As for retention, there are a few studies with humans on the directive influence of acquired drives (Rapaport, 1950), but none on primary drives.

The present study was undertaken to provide answers to the following questions: Does hunger at time of learning facilitate the learning and retention of words referring to food? Does hunger at time of recall facilitate the recall of words referring to food? Is there an interaction of hunger at time of learning and at time of recall upon recall of words referring to food? Are directive effects of hunger on learning and retention fairly general or do they vary as a function of specific characteristics of the stimulus?

A problem in working with the hunger drive in the laboratory is that if relatively high states of drive are to be obtained it is necessary to require abstinence from eating, which produces set effects that have been demonstrated to exert an influence as great as, or greater than, drive effects (Clarke & Epstein, 1957; Postman & Crutchfield, 1952; Taylor, 1956). Fortunately, several studies working within the normal food cycle have demonstrated directive effects of drive as a function of 4–6 hours of deprivation (Epstein & Smith, 1956; Lazarus, Yousem, & Arenberg, 1953; Postman & Crutchfield, 1952; Sanford, 1936, 1937).

## METHOD

*Subjects.* Sixty subjects, paid at the rate of $1.00 per hour, were enlisted from four fraternities at the University of Massachusetts. Subjects were randomly assigned to four equal groups according to a $2 \times 2$ factorial design of hunger at time of learning and hunger at time of memory. Fifteen subjects were hungry at learning but not hungry at recall, 15 were hungry at both learning and recall, 15 were not hungry at learning but were at recall, and 15 were neither hungry at learning nor recall. Hungry subjects were obtained by testing at 5:00 P.M., before the evening meal, and control subjects by testing at 6:30 P.M., after the evening meal. The group that was not hungry at learning and hungry at recall had a 23-hour memory interval, whereas the group that was hungry at learning and not hungry at recall had a 25-hour memory interval. Although the difference of 2 hours might be expected to influence total recall, there was no reason to believe it would influence selective recall, and, as it turned out, total recall was unaffected.

*Procedure.* For the learning session, a paired-associates list containing food words and control words was presented in four different random orders by a modification of a group method described by Saltz and Myers (1955). Standard instructions for learning paired associates were given, and subjects were required to raise their hands after reaching the criterion of one trial without error for the entire list. In order to reduce temptation to misrepresent, an appeal to honesty was made, and subjects were told they would have to remain for the entire session no matter when they were finished. Moreover, as each trial was represented by a separate page, cheating would have required turning pages, which could have been detected by the experimenter. In order to prevent overlearning for the more rapid learners, subjects were required to work on an anagrams task following completion of the learning task. From all observations it was evident that the instructions were adhered to. On the day following learning, three measures of memory were obtained. First the subject wrote down all the single words he recalled from the previous day, treating stimulus and response words independently. Next, the subject was given a test of recognition in

[1] This paper was presented, in part, at the American Psychological Association convention in 1958. The study was supported by Grant M-1293 from the National Institute of Mental Health, United States Public Health Service, as part of a project on the measurement of drive and conflict. At the time both authors were at the University of Massachusetts.

[2] Drive throughout this paper is defined as a state of activation with directive properties. For those who prefer to define drive as a state of activation only, the word "motive" could be substituted wherever the word drive appears.

which correct and incorrect words were included. Finally, a measure of relearning was obtained. The last two measures were discarded when it was found that scores were uniformly so high that there was not sufficient variance left to investigate.

*Paired-associates list.* The words to be learned consisted of 12 paired associates subdivided according to food relatedness and associative strength. High associative (HA) pairs consisted of a stimulus member paired with the word most often given to it in a preliminary word association test administered to 100 subjects from a comparable population. Low associative strength (LA) pairs consisted of a stimulus member paired with a word which it had never elicited as a response. There were four paired associates in which both members were food words—FRUIT-APPLE (HA), CHEESE-CRACKER (HA), WAFFLE-STEAK (LA), and CAKE-HAM (LA); four pairs in which both members were control words—CARPET-RUG (HA), LAMP-LIGHT (HA), ROOM-SOFA (LA), BED-WINDOW (LA); and four pairs in which one of the members was a food word and the other a control word, with order counterbalanced—EGG-STAIR (LA), STEW-CEILING (LA), HOUSE-BUTTER (LA), and SCREEN-POTATO (LA). The control words were related to home objects in order to partly control for the uniform concept formed by the food words. The pairs that contained words from both concepts, henceforth referred to as the "mixed" category, were necessarily of low associative strength, resulting in an overall design that was not balanced, but from which balanced designs could be extracted.

## RESULTS

*Learning.* Scores consisted of the first trial at which two successive correct anticipations occurred. The data were first treated by an analysis of variance in a three-dimensional design of hunger at time of learning, food relevance of paired associates, and associative strength of paired associates, each represented at two levels. The four paired associates involving mixed concepts were not included in this analysis as they did not occur at the level of high associative strength. In Table 1 it can be seen that hunger is significant as a main effect (.05 level). The hungry group learned words across classifications in fewer trials than the control group. However, there is no support for differential learning of pairs of food words and pairs of control words as a function of hunger (see sources of variance F × H and A × F × H). The remaining sources of variance, which are of only incidental interest, show that pairs of high associative strength were learned more rapidly than pairs of low associative strength (.01 level), and pairs of food words were learned more rapidly than pairs of control words (.01 level). There are significant differences (.01 level) between individual words within

TABLE 1

ANALYSIS OF VARIANCE OF TRIALS TO CRITERION AS A FUNCTION OF HUNGER, ASSOCIATIVE STRENGTH, AND FOOD RELEVANCE

| Source of variance | *df* | *MS* | *F* |
|---|---|---|---|
| Total between | 59 | | |
| Hunger (H) | 1 | 36.85 | 4.50* |
| Individual subjects (*Ss*) | 58 | 8.19[a] | |
| Total within | 420 | | |
| Associative strength (A) | 1 | 531.30 | 116.77** |
| A × H | 1 | 16.50 | 3.63 |
| A × *Ss* | 58 | 4.55[a] | |
| Food relevance (F) | 1 | 107.35 | 41.61** |
| F × H | 1 | .61 | — |
| F × *Ss* | 58 | 2.58[a] | |
| A × F | 1 | 55.36 | 21.13** |
| A × F × H | 1 | 2.84 | 1.08 |
| A × F × *Ss* | 58 | 2.62[a] | |
| Words within categories (W) | 4 | 23.14 | 8.67** |
| W × H | 4 | 4.93 | 1.85 |
| W × *Ss* | 232 | 2.67[a] | |
| Total | 479 | | |

[a] Error term for the mean squares listed above it up to the next error term.
* Significant at .05 level.
** Significant at .01 level.

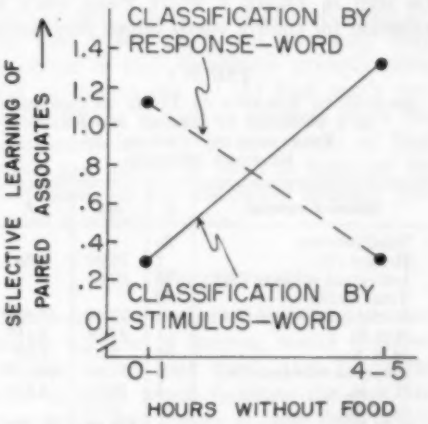



FIG. 1. Selective learning of food related paired associates as a function of whether the food word is a stimulus or response member. (Scores consist of mean trials to criterion for nonfood paired associates minus food paired associates, with food relevance determined independently for stimulus and response members. A high score indicates that food relevant paired associates were learned relatively rapidly.)

categories, and there is a significant interaction (.01 level) between associative strength and food relevance, with a greater disparity between the learning of pairs of food and control words of low associative strength than between the learning of food and control words of high associative strength.

In a second analysis of the learning data, the mixed pairs, which had not been previously included, were substituted for the pairs of high associative strength, thereby producing an orthogonal arrangement according to whether a word was a stimulus member or a response member and a food word or a control word. Thus, the following combinations of words were represented in the paired associates: food-food, food-control, control-food, and control-control. The analysis of variance was identical to the preceding analysis except that the classification by stimulus and response membership replaced the classification by low and high associative strength. It was found that the hungry group learned paired associates relatively more rapidly than the control group when the stimulus member was a food word and less rapidly when the response member was a food word. This can be seen in Figure 1 where mean trials to criterion for control words minus food words is plotted as a function of time without food. (By subtracting food words from control words, rather than the other way around, all values were kept positive, and a high score made to indicate relatively rapid learning of paired associates containing food words.) The mean differences for the hungry and control groups are, respectively, 0.29 and 1.37 for classification by stimulus word and 1.12 and 0.31 for classification by response word. The analysis of variance indicates that both effects are significant at the .05 level (see Table 2, sources of variance $H \times S_r$ and $H \times R_r$). However the significant interaction of hunger and words within categories (see source of variance $H \times W$) indicates that the findings may be a result of the particular words selected, and that verification with a different set of paired associates is required.

TABLE 2

ANALYSIS OF VARIANCE OF TRIALS TO CRITERION AS A FUNCTION OF HUNGER AND FOOD RELEVANCE OF STIMULUS AND RESPONSE MEMBERS

| Source of variance | *df* | *MS* | *F* |
|---|---|---|---|
| Total between | 59 | | |
| Hunger (H) | 1 | 76.40 | 2.56 |
| Individual subjects (Ss) | 58 | 29.81 | |
| Total within | 420 | | |
| Stimulus relevance ($S_r$) | 1 | 76.40 | 12.67** |
| $H \times S_r$ | 1 | 32.68 | 5.41* |
| $Ss \times S_r$ | 58 | 6.03[a] | 1.29 |
| Response relevance ($R_r$) | 1 | 57.86 | 14.57** |
| $H \times R_r$ | 1 | 18.48 | 4.65* |
| $Ss \times R_r$ | 58 | 3.97[a] | |
| $S_r \times R_r$ | 1 | 129.64 | 21.72** |
| $H \times S_r \times R_r$ | 1 | 1.40 | — |
| $Ss \times S_r \times R_r$ | 58 | 5.97[a] | 1.28 |
| Individual words (W) | 4 | 39.41 | 33.83 |
| $H \times W$ | 4 | 23.73 | 5.09** |
| $Ss \times W$ | 232 | 4.66[a] | |
| Total | 479 | | |

[a] Error term for the mean squares listed above it up to the next error term.
* Significant at .05 level.
** Significant at .01 level.

TABLE 3

ANALYSIS OF VARIANCE OF NUMBER OF WORDS RECALLED AS A FUNCTION OF HUNGER AT TIME OF LEARNING, HUNGER AT TIME OF RECALL, ASSOCIATIVE STRENGTH, AND FOOD RELEVANCE

| Source of variance | *df* | *MS* | *F* |
|---|---|---|---|
| Total between | 55 | | |
| Hunger at learning ($H_l$) | 1 | .76 | — |
| Hunger at recall ($H_r$) | 1 | .10 | — |
| $H_l \times H_r$ | 1 | 1.45 | 1.02 |
| Individual subjects (*Ss*) | 52 | 1.41[a] | |
| Total within | 280 | | |
| Associative strength (A) | 2 | 16.75 | 12.50** |
| $A \times H_l$ | 2 | .08 | — |
| $A \times H_r$ | 2 | .25 | — |
| $A \times H_l \times H_r$ | 2 | 3.18 | 2.37 |
| $A \times Ss$ | 104 | 1.34[a] | |
| Food relevance (F) | 1 | .29 | |
| $F \times H_l$ | 1 | .30 | — |
| $F \times H_r$ | 1 | 2.35 | 2.67 |
| $F \times H_l \times H_r$ | 1 | 8.03 | 9.12** |
| $F \times Ss$ | 52 | .88[a] | |
| $A \times F$ | 2 | 11.76 | 8.98** |
| $A \times F \times H_l$ | 2 | 4.15 | 3.17* |
| $A \times F \times H_r$ | 2 | 2.36 | 1.80 |
| $A \times F \times H_l \times H_r$ | 2 | .95 | — |
| $A \times F \times Ss$ | 104 | 1.31[a] | |
| Total | 335 | | |

[a] Error term for the mean squares listed above it up to the next error term.
* Significant at .05 level.
** Significant at .01 level.

*Recall.* Recall was analyzed by an analysis of variance for the variables of hunger at learning, hunger at recall, food relevance of the word recalled, and associative strength, with the mixed category treated as an additional level of associative strength. Scores

were the number of words recalled within each of the six classifications formed by the two levels of food relevance and the three levels of associative strength. There were 14 entries rather than 16 for each condition, as 2 subjects failed to appear for the recall session.

In Table 3 it can be seen that there is a significant interaction (.01 level) for the variables of hunger at learning, hunger at recall, and food relevance of the word recalled. In Figure 2 it is apparent that the reason for the interaction is that an increase in recall of food words as a function of hunger at recall occurs only when the groups not hungry at learning are compared. If only subjects not hungry at learning are considered, the mean number of food words minus control words recalled is −1.79 for the condition of not hungry at recall, and 0.93 for the condition of hungry at recall. In contrast, when the same comparison is made for subjects who learned when hungry, the relative recall of food words varies little as a function of hunger at recall, with mean differences, respectively, of 0.43 and −0.29 for the not hungry and hungry conditions. It is of interest that neither hunger at learning nor at recall approaches significance for total words recalled. Apparently, hunger decreases recall of control words to about the same extent that it increases recall of food words.

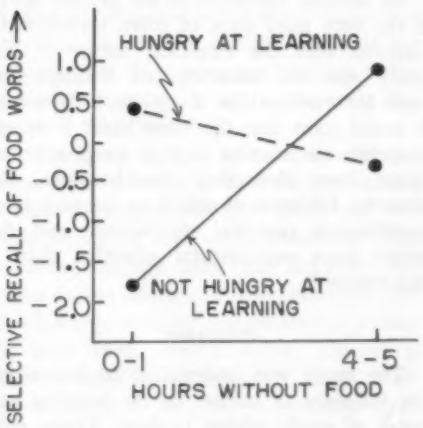



FIG. 2. Selective recall of food words as a function of hunger state during learning and during recall. (Scores consist of number of food words minus nonfood words recalled.)

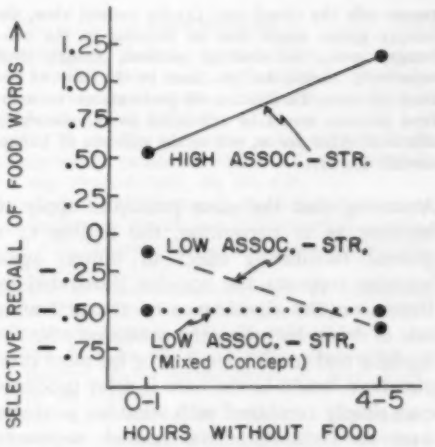



FIG. 3. Selective recall of food words as a function of hunger at learning and associative strength of the paired associates presented in the learning session. (Scores consist of number of food words minus nonfood words recalled.)

There is a significant interaction for the variables of hunger at learning, food relevance, and associative strength (.05 level). In Figure 3 it can be seen that the relationship is a direct one for hunger at time of learning and relative recall of food words for the food words which had previously been presented in high associative strength pairs, and an inverse one for the food words presented in low associative strength pairs. No relationship between relative recall of food words and hunger at recall is indicated for food words of low associative strength presented in mixed concepts.

## DISCUSSION

It would seem reasonable to expect that hunger at time of learning would facilitate the learning of food words, that hunger at time of recall would facilitate the recall of food words, and hunger at both times would result in the greatest degree of selective recall. Brown (1953), in a highly relevant discussion of the influence of hunger upon perception states that,

> it would not be suprising if the hungry *Ss*, who possess both the internal cues and the visual cues, should do better than the nonhungry *Ss*, who

> posses only the visual cues. On the present view, the hungry group might also be superior to the nonhungry group on nonfood pictures, though their superiority should be less than in the case of the food pictures. Facilitation of performance on nonfood pictures would be attributed to the energizing effects of drive *per se*, not to the presence of hunger stimuli (p. 9).

Assuming that the same principles apply to learning as to perception, the finding of a general facilitating effect of hunger upon learning supports the position illustrated by Brown on the significance of the activating role of drive. However, the results on selective learning and recall appear to be far more complex than would be the case if drive produced cues simply combined with stimulus produced cues to facilitate drive related responses. Several studies on the influence of hunger upon perception and imagination have suggested that associated with every drive state there is a drive oriented autistic process and a reality oriented inhibitory process (Epstein, 1961; Epstein & Smith, 1956; Levine, Chein, & Murphy, 1942; McClelland & Atkinson, 1948; Sanford, 1937; Wispé, 1954). It is possible that the finding of facilitation in learning when the stimulus member is a food word was an example of the autistic process, and that the retardation in learning when the response member was a food word of the inhibitory process. If this is so, it would suggest that drives tend to induce sensitization in perception by inducing set effects or by some other means, and to induce defensiveness in overt response.

In regard to recall, it was found that the directive effect of hunger upon recall depended upon the hunger condition at learning. Only when learning took place in the nonhungry condition was selective recall as a function of hunger at recall demonstrated. This would suggest that drive produced cues can interfere with as well as facilitate recall of drive related material.

The finding that hunger increased recall of food words without influencing total recall suggests that the directive effects of a drive may function as an attention or set producing phenomenon, i.e., by focusing attention in one direction, attention in other areas is diminished. In this respect, drive produced set effects differ from other set effects only to the extent that the cues are internal and therefore less likely to be correctly labeled.

The effect of hunger on selective learning was more equivocal than its effect on selective recall. This is not surprising, as recall involves a greater separation from the stimulus than does learning. An increase in memory interval can be viewed in the same manner as an increase in ambiguity, i.e., both factors by reducing the degree to which the response is determined by the stimulus, increase the degree to which motivational and incidental factors play a role.

The finding that hunger during learning resulted in an increase in selective recall of food words of high associative strength but a decrease in recall of food words of low associative strength can be understood by considering the particular combinations of words investigated. High associative strength pairs consisted of combinations, such as CHEESE-CRACKER, which are commonly experienced together during reduction of hunger, while the same does not apply to the low associative strength combinations, such as WAFFLE-STEAK. It is likely that a drive arouses only those associations that have previously been associated with it or with its reduction, rather than all stimuli related to the drive in the individual's cognitive framework.

An obvious limitation in the present study is the very mild state of drive investigated. Coupled with the empirical nature of the study and the unanticipated findings, the need for verifications is apparent. However, it would seem that the "new look" is as appropriate for learning as it is for perception. Apart from theoretical considerations, the directive influence of drives on learning is of considerable practical significance and deserves more experimental attention than it has received.

## SUMMARY

The study was undertaken to investigate the influence of hunger on the learning and recall of words related to food. Thirty subjects learned a list of paired associates consisting of all possible combinations of food and nonfood words, immediately before the

evening meal. Thirty other subjects learned the same paired associates immediately after the evening meal. Half of each group was tested for recall the next day before the evening meal, and half after the evening meal. The recall task consisted of writing all the single words that could be remembered from the learning session of the previous day.

The major findings and conclusions were as follows:

1. The group hungry at learning learned more rapidly than the control group.

2. The group hungry at learning demonstrated relative facilitation in the learning of paired associates when the stimulus word was a food word, and relative retardation when the response word was a food word. This was interpreted as indicating that drives tend to induce selective facilitation in perception but inhibition in response for drive related material.

3. There was an interaction of hunger condition at learning and at recall upon selective recall of food words. Only when comparison was between the groups who were not hungry during learing was hunger during recall found to be directly associated with selective recall of food words. It was concluded that drive produced cues interfere with as well as facilitate recall of drive related material.

4. Increased recall of food words as a function of hunger occurred at the expense of a decrease in recall of control words. It was suggested that drive creates a selective set for drive related materials and consequent inattention to other material.

5. The group hungry at learning exhibited greater selective recall of food words presented in high associative strength pairs and less selective recall of food words presented in low associative strength pairs than the group not hungry during learning. It was concluded that drive tends to activate associations of stimuli that have been present during arousal or reduction of the drive, rather than of all stimuli related to the drive in the individual's cognitive framework.

## REFERENCES


BROWN, J. S. Problems presented by the concept of acquired drives. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1953.* Lincoln: Univer. Nebraska Press, 1953. Pp. 1–21.

CLARK, R. A. The projective measurement of experimentally induced levels of sexual motivation. *J. exp. Psychol.*, 1952, **44**, 391–399.

CLARK, A. R., & EPSTEIN, S. Food-related responses to ambiguous stimuli as a function of time without food and experimental set. *Amer. Psychologist*, 1957, **12**, 394. (Abstract)

EPSTEIN, S. Food-related responses to ambiguous stimuli as a function of hunger and ego strength. *J. consult. Psychol.*, 1961, **25**, 463–469.

EPSTEIN, S., & SMITH, R. Thematic apperception as a measure of the hunger drive. *J. proj. Tech.*, 1956, **20**, 373–384.

EPSTEIN, S., & SMITH, R. Thematic apperception, Rorschach content, and ratings of sexual attractiveness of women as measures of the sex drive. *J. consult. Psychol.*, 1957, **21**, 473–478.

LAZARUS, R. S., YOUSEM, H., & ARENBERG, A. Hunger and perception. *J. Pers.* 1953, **21**, 312–328.

LEVINE, R., CHEIN, I., & MURPHY, G. The relation of the intensity of a need to the amount of perceptual distortion. *J. Psychol.*, 1942, **13**, 283–293.

MCCLELLAND, D. C., & ATKINSON, J. W. The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *J. Psychol.*, 1948, **25**, 205–222.

MURRAY, E. J. Conflict and repression during sleep deprivation. *J. abnorm. soc. Psychol.*, 1959, **59**, 95–101.

POSTMAN, L., & CRUTCHFIELD, R. S. The interaction of need, set, and stimulus-structure in a cognitive task. *Amer. J. Psychol.*, 1952, **65**, 196–217.

RAPAPORT, D. *Emotions and memory.* (2nd ed.) New York: International Univer. Press, 1950.

SALTZ, E., & MYERS, T. I. A method for group presentation of paired-associate learning materials. Paper read at Midwestern Psychological Association convention, Chicago, 1955.

SANFORD, R. N. The effect of abstinence from food upon imaginal processes: A preliminary experiment. *J. Psychol.*, 1936, **2**, 129–136.

SANFORD, R. N. The effect of abstinence from food upon imaginal processes: A further experiment. *J. Psychol.*, 1937, **3**, 145–159.

TAYLOR, JANET A. Physiological need, set, and visual duration threshold. *J. abnorm. soc. Psychol.*, 1956, **52**, 96–99.

WISPÉ, L. G. Physiological need, verbal frequency, and word association. *J. abnorm. soc. Psychol.*, 1954, **49**, 229–234.


(Received February 2, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 136–142

# VALUES EXPRESSED IN AMERICAN CHILDREN'S READERS: 1800–1950 [1]

RICHARD DECHARMS AND GERALD H. MOELLER

*Washington University*

Students of cultural change within the United States seem to have reached some agreement as to a trend observable within the last century. This trend, which deals with some of the basic values of our culture, may be seen as a change from what Weber (1930) called "the Protestant ethic" to what has been called the "social ethic" (Whyte, 1956). Specifically, the dominant value of individual salvation through hard work, thrift, and competition is seen as being replaced by "a belief in the group as the source of creativity; a belief in 'belongingness' as the ultimate need of the individual; and a belief in the application of science to achieve the belongingness" (Whyte, 1956, p. 7). In Riesman's (Riesman, Glazer, & Denney, 1950) terminology the basic trend is from inner-direction to other-direction.

Actually this process is circular in the sense that the cultural change is probably accompanied by a change in values which starts a new cycle. The psychologist likes to conceive of the basic variables in human behavior as being internally determined and thus breaks this circle and concentrates on motives as basic.

The aim of the present paper is to investigate psychological variables which it seems logical to predict will be associated with the cultural changes observed in the United States over the last century and a half.

McClelland (1955) has noted striking similarity between his concept of the person with high achievement motivation and Riesman's inner-directed character type. McClelland defines achievement motivation as "success in competition with a standard of excellence" (p. 43). According to Riesman et al. (1950) "the drive instilled in the child is to *live up to ideals* and to test his ability to be on his own by continuous experiments in self-mastery" (p. 59). For the other-directed person "making good becomes almost equivalent to making friends" (Riesman et al., 1950, p. 66). This sounds very much like the person with high affiliation motivation. Atkinson (1958) defines affiliation motivation as "concern . . . over establishing, maintaining, or restoring a positive affective relationship with another person. This relationship is most adequately described by the word friendship" (p. 206).

Assuming that achievement motivation is a basic component of the inner-directed character type, and that affiliation motivation is a basic component of the other-directed character type, in the context of Riesman's cultural change thesis, one would predict a decline in over-all achievement motivation and an increase in affiliation motivation in the last century of United States history. Strauss and Houghton (1960) have found evidence giving some support to these hypotheses in the period since 1924 in a study of 4-H club journals. A meaningful relationship should also be found between achievement orientation and economic and technological change according to McClelland (1955).

A further aspect of Riesman's thesis is that the stage of inner-direction is preceded by a stage of tradition-direction. During this stage strict moral codes demand behavioral conformity of the individual. The change from a tradition-directed society to an inner-directed society involves a secularization in the sense that the individual must prove himself worthy as in Weber's (1930) Protestant ethic, rather than be told what to do by categorical imperatives. One might thus predict more reliance on moral teaching early in the history of the United States.

[1] This study was carried out under Contract Nonr-816 (11) between the Office of Naval Research and the Small Groups Laboratory, Social Science Institute at Washington University.

We wish to thank David C. McClelland and Raymond E. Callahan for critical comments on an earlier version of the paper.

The objective measurement of cultural orientations or values is always difficult, especially if an attempt is to be made to tap the values of the past. An intriguing attempt to measure the motives of an ancient culture has been reported by McClelland (1958). He has developed a method of assessing achievement motivation in a culture by content analysis of literary products of the culture. Using this tool he found striking confirmation for his hypothesis that achievement motivation preceded the economic and technological development of Athenian civilization in classical Greece, a culture also discussed by Riesman.

McClelland's measure of motivation was developed originally to assess the motives of individuals (McClelland, Atkinson, Clark, & Lowell, 1953). The extension of it to apply to cultures raises questions as to what is being measured. McClelland (1958) has argued that his content analysis of documents produced by the culture will give a measure of the level of specific types of motivation within the culture if the documents are carefully selected to reduce effects of other obviously important variables. The best support for this argument derives from a study which demonstrated that content analysis of American Indian folk tales relates to content analysis of stories obtained from individuals in the manner used originally to validate the measure for studying individual motives (McClelland & Friedman, 1952).

It is obvious that a measure which has been shown to be related to individual motives would be expected to reflect the motives of the writer of any document chosen from a culture. In order to use such a measure as an indication of cultural orientation one does not have to assume that the motive score of an individual author is a measure of the cultural orientation alone. One must, however, assume that a portion of the score is a measure of the cultural orientation. The problem of reducing idiosyncratic components in the measure of cultural orientation becomes one of (*a*) sampling randomly many authors (*b*) under as similar conditions as possible (i.e., writing similar material) and (*c*) choosing materials which should place few restrictions on the author's fantasy.

All of the above advantages can be obtained by careful sampling of stories written for children's readers. In addition, the stories are actually written to be used in transmitting cultural values, and information on how widely they have been used gives at once some indication of cultural acceptance of the values contained in the book and of the extent of its influence. An example of the use of children's readers to assess values in many cultures has been presented by McClelland (1961).

The schema presented above predicts a relationship between cultural achievement orientation and behavior of the members of the culture which would lead to technological advance. Just as one might predict that an individual with high achievement motivation might strive for some unique accomplishment, one might also predict that a culture with strong achievement orientation would produce many inventions. A measure of the inventiveness of the culture at various periods in history might be obtained from the number of patents issued per population, and one could predict a relationship between this and a measure of cultural achievement orientation.

### *Hypotheses*

The present study is an attempt to plot the incidence of achievement and affiliation imagery and moral teaching in a sample of children's readers from 1800 to the present. In addition, achievement imagery is to be related to data on the number of patents issued per population.

*Hypothesis I.* The incidence of achievement imagery in a sample of children's readers selected over the period 1800–1950 will decrease over the time period.

*Hypothesis II.* The incidence of affiliation imagery in the same sample of readers will increase over the time period.

*Hypothesis III.* The incidence of moral teaching in the sample will decrease over the time period.

*Hypothesis IV.* The incidence of achievement imagery will be positively related to the number of patents issued, corrected for changes in population.

The hypotheses assume, with Riesman, that the nineteenth century in the United States was dominated by the inner-directed charac-

ter type. Riesman is not specific as to dates, but it would appear that the early period of the century witnessed the transition from the tradition-directed character type and that the United States has recently been in transition from an inner-directed phase to an other-directed phase.

## METHOD

A bibliography of reading textbooks with copyright dates ranging from 1800 to 1952 was compiled. An attempt was made to procure at least four books from each 20-year period beginning in 1800. Readers were excluded which were not in wide use [2] during the period or which were used by religious affiliated schools. In the more recent periods from which more than four books were available, the choice of books was made randomly. In the periods in which fewer than four books were available the sample from each book was enlarged in so far as possible. Generally, the sample from each book was obtained by scoring every third page.

It was found that the number of words per page was sufficiently similar throughout the total sample to allow use of the page itself as the scoring unit. In order to equate for number of pages available the score was the number of pages containing imagery per 75 pages sampled. A raw score was thus computed for every 75 pages sampled (i.e., 25 pages scored).

The readers chosen for the study were, generally speaking, at a fourth grade level. During the nineteenth century many readers were designated in ways which had no relevance to grade level or, as in some cases, grade level was quite different from that of contemporary American readers in which the vocabulary is based on standard word lists. In some instances it was necessary to use the Dale and Chall (1948) formula [3] for predicting readability to determine whether the readers might be allowed in the study.

The pages selected from each book were scored independently by two scorers as to whether the page contained (*a*) achievement imagery, (*b*) affiliation imagery, or (*c*) a category called moral teaching. Achievement and affiliation imagery were scored according to the procedure outlined in Atkinson (1958). The subcategories usually scored in this procedure were not scored. The category Moral Teaching was developed and defined as explicit or implicit statements of judgment between right and wrong from the point of view of the author. The following (McGuffey, 1857) are examples of items which were scored for moral teaching:

> The little boy took care of his faithful dog as long as he lived and never forgot that we must do good to others, if we wish them to do the same to us (p. 42).
>
> Now that is the way with a great many thoughtless, quick tempered people. They try to find fault with somebody or something else, and get into a passion, and perhaps do mischief, when, if they would but reflect a little, it is their own dear selves who ought to bear the blame (p. 47).

Scorer reliabilities, based on presence of imagery only, were consistently high (Achievement Imagery = 94%, Affiliation Imagery = 96%, Moral Teaching = 97%).

The number of patents issued by the United States Patent Office and the United States Census figures were taken from governmental documents (United States Department of Commerce, 1960) and a patent index was computed by dividing the number of patents granted in a 20-year period by the population reported in the midyear of that period and multiplying by one million. This results in an index of patents issued per one million population during the period.

There are two methodological flaws in the procedure which it was felt might have had an effect on the results. In the first place, the technique of blind scoring was not employed. The scoring was done directly from the book and it was therefore probable that the scorer knew the date of the book. The effect of this knowledge cannot be assessed. A second methodological flaw lies in the sampling procedure. Systematic samples were taken from each book and the books were chosen as representative and in wide use. However, since each score was based on 75 pages of text, some books were more heavily weighted than others and the individual values of their authors might have unduly influenced the results.

In order to correct these methodological flaws it was decided to repeat the study with a drastically smaller sample. A sample of 5 pages was chosen at random from each book. The sampling of books followed the same criteria as those used in the first sample. Four books were selected from each period except the periods 1800–1819 and 1820–1839 where only two were available. The books from these two periods were double sampled. Scores on each variable were assigned to each book giving four scores for each of the eight periods thus resolving the ambiguities of sampling in the first study.

The sample had to be drastically cut since the

[2] Evidence of "wide use" was fairly easy to establish in readers published after 1850 since the McGuffey readers and the readers of the major book companies enjoyed national popularity as official texts of large school systems. Prior to 1850 Johnson (1904) and Nietz (1961) provided lists of historical texts from which to choose. In the first two decades of the sample the only secular texts available were used.

The following criteria were used in this order to establish "wide use": Evidence (often cited in later editions) of number of copies sold; several editions of the same book; and knowledge of use by large school systems.

[3] This readability formula which utilizes a word list and sentence length for determining grade level of reading materials was used because of evidence of its applicability to middle grade reading matter and high correlations with other formulae and criteria of readability (Chall, 1958).

pages were typed and coded for blind scoring. (A total of 192 pages were scored independently by two scorers in the replication whereas 2,375 pages were scored in the first study.) The same number of typed lines were taken starting with each page which had been drawn randomly.

Whenever available different books were selected for the second sample. It was anticipated that since this was a much smaller sample than the first the results would not be as statistically significant. It was felt that general trends in the same direction would validate the statistically significant findings of the first sample. Actually, plots of the results of the two studies are almost identical and statistical significance was reached in most instances in the second study, although, as anticipated, the probability levels were not as great as in the first. This comparison of the two replications gives greater confidence in the results of the first study. The data presented here come from the first study. Statistical analyses will be presented for both studies.

## RESULTS

Table 1 presents the mean imagery scores for achievement, affiliation, and moral teaching in each of the 20-year periods. Figures 1 and 2 are graphic presentations of these data.

Hypothesis I predicted a consistent decrease in achievement imagery. The data (see Table 1 and Figure 1) show a sharp decline since 1890, but a steady increase from 1800 to the peak at about 1890. The second sample showed almost an identical curve with consistent increase up to about 1890 and then a sharp decline. The data of both samples show a significant relationship between amount of imagery and date (First sample, $F = 8.09$, $df = 7/87$, $p < .0005$; Second sample, $F = 2.62$, $df = 7/24$, $p < .05$).

TABLE 1

MEAN SCORES FOR ACHIEVEMENT IMAGERY, AFFILIATION IMAGERY, AND MORAL TEACHING IN THE FIRST SAMPLE OF READERS

| 20-year midpoint | $N$[a] | Achievement imagery | | Affiliation imagery[b] | | Moral teaching[b] | |
|---|---|---|---|---|---|---|---|
| | | $\bar{X}$ | $\sigma^2$ | $\bar{X}$ | $\sigma^2$ | $\bar{X}$ | $\sigma^2$ |
| 1810 | 3 | 2.67 | 4.5 | 3.33 | 2.5 | 16.00 | 22.0 |
| 1830 | 4 | 2.50 | 1.7 | 4.25 | 11.7 | 16.75 | 9.7 |
| 1850 | 12 | 4.42 | 13.4 | 6.00 | 24.4 | 12.42 | 11.2 |
| 1870 | 3 | 8.33 | 2.5 | 6.33 | 25.5 | 6.00 | 1.0 |
| 1890 | 16 | 11.06 | 13.5 | 5.13 | 4.7 | 4.19 | 4.9 |
| 1910 | 10 | 9.40 | 5.1 | 6.70 | 12.0 | 4.50 | 2.7 |
| 1930 | 15 | 6.33 | 19.6 | 9.33 | 6.1 | 1.00 | 0.7 |
| 1950 | 32 | 4.25 | 14.8 | 5.50 | 12.7 | 0.06 | 0.1 |

[a] Number of samples of 25 pages scored. The raw score was the number of pages (out of 25) containing imagery.
[b] Variance heterogeneous. In no case were the variances heterogeneous in the second sample.

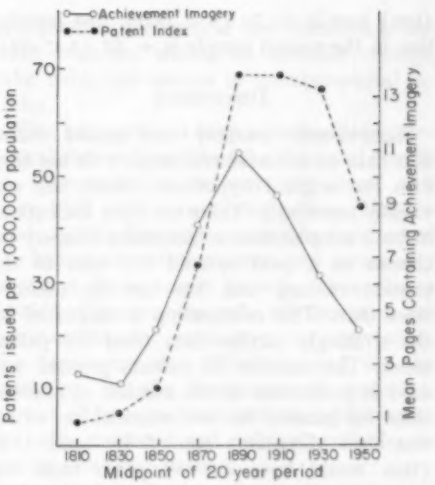



FIG. 1. Mean number of pages (out of 25) containing achievement imagery and the patent index.

Hypothesis II predicted a consistent increase in affiliation imagery from 1800 to the present. The data in Table 1 and Figure 2 show no consistent increase but do show a general trend with an unexpected drop in 1950 to the 1890 level. Analysis of variance shows a significant relationship ($F = 2.41$, $df = 7/87$, $p < .05$) but the variance was heterogeneous. The Kruskal-Wallis (see Siegal, 1956) analysis of variance of ranks also shows a significant relationship ($p < .01$). In addition, a rank correlation (Kendall's tau) between affiliation imagery and date equals + .61 ($p < .03$).

The data from the second sample show the same general trend with greater variability. The correlation drops below significance and analysis of variance does not reach significance.

Hypothesis III predicts a decrease in moral teaching from 1800 to the present. The data show a striking confirmation (First sample, $F = 101.9$, $df = 7/87$, $p < .0005$, heterogeneous variance; Second sample, $F = 6.95$, $df = 7/24$, $p < .0005$, variance homogeneous).

Hypothesis IV predicts a relationship between the amount of achievement imagery during a specific period and the index of patents issued per population. Figure 1 shows this relationship in graphic form. A striking relationship is apparent. The rank correlation

(tau) here is + .79 ($p < .003$). The correlation in the second sample is + .68 ($p < .02$).

## DISCUSSION

*Achievement imagery and patent index.* The data on achievement imagery do not confirm the original hypothesis which was obviously too simple. There are clear indications in both samples that achievement imagery increases to a peak around the turn of the present century and has steadily declined since then. This relationship is supported by the strikingly similar data from the patent index. The number of patents granted was used in preference to the number of applications for patents for two reasons: no record was kept of patents applied for until 1840 (this would have cut 39 years from the patent/population measure) and the very fact that a patent is issued is indicative of the "uniqueness" of the patent. Unique accomplishment is one of the criteria for scoring achievement imagery (Atkinson, 1958).

*Affiliation imagery.* The data from both samples tend to confirm the hypothesis of increasing affiliation imagery, although the results were not statistically significant in the second sample. Certain aspects of the difference in type of affiliation imagery through time are noteworthy. Much of the early and middle nineteenth century readers' affiliation imagery, though widely scattered, was quite unsophisticated as shown by the following (Parker & Watson, 1857) example:

> I love my dear little brother and I am pleased when I see him happy. I did not intend to disobey you, dear father, and I hope you will not be displeased with me for what I have done (p. 75).

The discussion of the joys of giving to others continued unabated for the next page and a half. In contrast, affiliation imagery in the period from 1920 was considerably more subtle.

It was more difficult to score affiliation imagery in the earliest books in the study and it was in this period that interrater reliability was lowest. The difficulty lay in differentiating true affective affiliation imagery from a culturally sanctioned form of address. Thus "dear son," "dear father," and "my dear" were not scored unless affect was also demonstrated since this was often mere conventionally approved formalism.

It is possible that this scoring difficulty could account for the results found. This, in combination with the fact that the results were not significant in the second sample, suggests caution in interpretation of these results.

*Moral imagery.* The decline in the religious-moral emphasis in textbooks has long been noted by various researchers. Hart (1933), in analyzing selected popular magazines of the period from 1900 to 1930, found evidence of a general decline in the status of religion and religious sanctions. These findings are in general accord with the results of the present study. In the case of the school readers it may be argued that the diminishing frequency of moral references is a result of the secularization of the schools during the nineteenth century.

It should be noted that the first schools in the colonies were church-sponsored and, in many instances, the minister of the church also served as teacher to the children. His primary purpose was not general education but simply to teach the pupils to read the Bible. The shift from church-sponsored to the

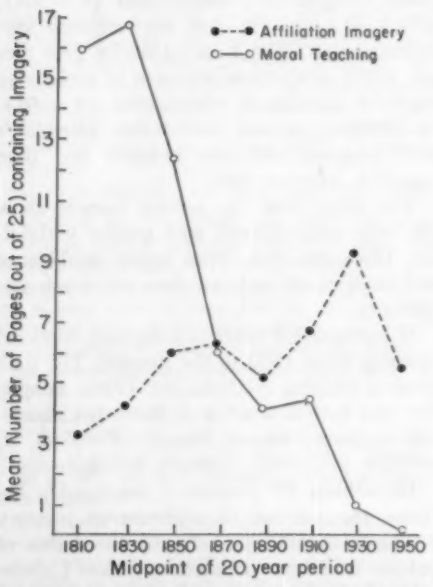



FIG. 2. Mean number of pages (out of 25) containing affiliation imagery and moral teaching imagery.

public-sponsored and supported schools began about the time of the American Revolution and continued through most of the nineteenth century. However, since all the books in the present study were prepared for public schools, the decrease in moral teaching imagery indicates the cultural trend toward secularization which affected the management of the schools and was reflected in the books written for the schools even after they had become nominally secular.

The antecedent conditions of changes in values such as demonstrated here are very complex. As noted earlier theorists such as Riesman et al. (1950) stress the importance of psychological factors (character type) which lead to the examination of child rearing practices. Economic historians stress political and economic factors (Rostow, 1960). The historical evidence is probably easier to marshal to support the economic interpretation.

The findings of the present study fit very well the conceptual paradigm of Rostow (1960) who stresses the economic factors. He has developed a general model of the stages of economic growth which distinguishes (*a*) precondition for take-off, (*b*) the take-off, (*c*) the drive to maturity, and (*d*) the age of high mass consumption.

In the United States, Rostow (1960) found that the traditional or agricultural society lasted until about 1840. The take-off occurred from 1843 to 1860. This appears to be the take-off period for achievement orientation also (see Figure 1). During the next period from 1860 to 1900, called by Rostow the drive to maturity, "some ten to twenty per cent of the national income was steadily invested, permitting industrial output regularly to outstrip the increase in population" (p. 9). The United States, according to Rostow's data and reasoning, reached technological maturity around 1900. This date is extremely close to the high points of achievement imagery and patent measures.

Rostow's (1960) preconditions for take-off are technological developments which might set the stage for increased social mobility, a factor mentioned by Riesman et al. (1950) as affecting child rearing practices. Rostow feels that during this period the idea that economic progress is possible and necessary for such ends as national dignity becomes prevalent in the culture and men come forward who are willing to mobilize savings, take risks, and engage in entrepreneurial activity.

The latter aspects have a distinctly psychological flavor. Men who take risks and engage in entrepreneurial activity are those who have high achievement motivation (McClelland, 1958). Recently economists have noted the importance of motivation and personality structure in economic growth. Thus Hagen (1958) discusses the role of the need for achievement, for autonomy, for aggression, for dominance, for affiliation, and for dependence in the beginning of economic growth. These motivational variables interact with economic and political variables to produce cultural changes.

As noted in the introduction, evidence for the importance of psychological factors such as motives and values in cultural change and economic growth has been presented by McClelland (1955, 1958, 1961). There are, moreover, studies which have made a start in uncovering the relationship between child rearing practices and achievement and affiliation motivation. Briefly, achievement motivation appears to be associated with early parental stress on independence training and mastery, coupled with a warm acceptance of the child (Winterbottom, 1958). Affiliation motivation is related to maternal acceptance and to parental stress on interpersonal involvement of the child (Gall, 1960).

These findings appear to be in accord with Riesman's analysis of the child rearing practices which lead to the inner- and other-directed character types. The parent rearing a child in the period of transition to inner-direction must equip him with a "gyroscope" which will fit him to remain on course in a society where it is impossible to foretell, due to increasing social mobility, what role he will be called upon to play. He must be equipped to be self-reliant and independent. These are the aspects which Riesman sees in nineteenth century child rearing. The antecedents of achievement motivation seem clear.

On the other hand, with increasing urbanization and population density which result from technological advance, the child is no longer pushed to be independent, but learns the importance of other individuals in the

environment. He must be taught to win approval. Although Riesman's argument is more complicated than this, the child rearing practices which he sees in contemporary United States culture seem to be ones which might lead to affiliation motivation.

In summary we propose that motivation, or cultural orientation, be conceived of as an intervening variable standing between antecedent environmental factors associated with economic and political changes and consequent behavior resulting in cultural changes such as technological growth. Such an analysis should give increased explanatory power, since it is probable that motivation is a function of factors other than economic changes. For instance, cultural values affect child rearing practices and hence motives (McClelland, Rindlisbacher, & deCharms, 1955). Thus two cultures undergoing similar economic or political change may react quite differently due to the intervening variables of values, child rearing practices, and motives.

## SUMMARY

Content analysis of children's readers from schools in the United States demonstrated a rise in achievement imagery from 1800 to about 1900 and then a steady decline. The achievement imagery curve was related to an index of the number of patents issued. A steady decline over the period 1800–1950 was found in the amount of moral teaching in readers. There was tentative indication of an increase in affiliation imagery.

The over-all picture presented by the data correspond very well with certain cultural trends pointed out by students of cultural change. The data illustrate an interesting technique for obtaining objective data to investigate cultural historical hypotheses.

## REFERENCES

ATKINSON, J. W. *Motives in fantasy, action, and society*. Princeton: Van Nostrand, 1958.

CHALL, JEANNE. Readability: An appraisal of research and application. *Bur. educ. Res. Monogr. Ohio St. U.*, 1958, No. 34.

DALE, E., & CHALL, JEANNE. A formula for predicting readability. *Educ. res. Bull.*, 1948, **27**, 11–20, 37–54.

GALL, H. S. The development of affiliation motivation. Unpublished doctoral dissertation, University of North Carolina, 1960.

HAGEN, E. E. How economic growth begins: A general theory applied to Japan. *Publ. opin. Quart.*, 1958, **12**, 373–390.

HART, H. Changing social attitudes and interests. In, *Recent social trends in the United States*. New York: McGraw-Hill, 1933. Pp. 382–442.

JOHNSON, C. *Old-time schools and school-books*. New York: Macmillan, 1904.

MCCLELLAND, D. C. Some social consequenes of achievement motivation. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1955*. Lincoln: Univer. Nebraska Press, 1955. Pp. 41–65.

MCCLELLAND, D. C. The use of measures of human motivation in the study of society. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society*. Princeton: Van Nostrand, 1958. Pp. 518–552.

MCCLELLAND, D. C. *The achieving society*. Princeton: Van Nostrand, 1961.

MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Croft, 1953.

MCCLELLAND, D. C., & FRIEDMAN, G. A. A cross-cultural study of the relationship between child-rearing practices and achievement motivation appearing in folk tales. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology*. (2nd ed.) New York: Holt, 1952. Pp. 243–249.

MCCLELLAND, D. C., RINDLISBACHER, A., & DECHARMS, R. Religious and other sources of parental attitudes toward independence training. In D. C. McClelland (Ed.), *Studies in motivation*. New York: Appleton-Century-Croft, 1955. Pp. 389–397.

MCGUFFEY, W. H. *New fourth eclectic reader*. New York: Wilson, Hinkle, 1857.

NIETZ, J. A. *Old textbooks*. Pittsburgh: Univer. Pittsburgh Press, 1961.

PARKER, R. G., & WATSON, J. M. *The national second reader*. New York: Barnes, 1857.

RIESMAN, D., GLAZER, N., & DENNEY, R. *The lonely crowd*. New Haven: Yale Univer. Press, 1950.

ROSTOW, W. W. *The stages of economic growth*. Cambridge: Cambridge Univer. Press, 1960.

SIEGAL, S. S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

STRAUSS, M. A., & HOUGHTON, L. J. Achievement, affiliation, and cooperation values as clues to trends in American rural society, 1924-1958. *Rural Sociol.*, 1960, **25**, 394–403.

UNITED STATES DEPARTMENT OF COMMERCE. *Historical statistics of the United States colonial times to 1957*. Washington: 1960.

WEBER, M. *The Protestant ethic*. (Trans. by T. Parsons) New York: Scribner's, 1930.

WHYTE, W. H. *The organization man*. Garden City: Doubleday, 1956.

WINTERBOTTOM, M. R. The relation of need for achievement to learning experiences in independence and mastery. In J. W. Atkinson, (Ed), *Motives in fantasy, action, and society*. New York: Van Nostrand, 1958. Pp. 453–478.

(Received February 6, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 143–147

# THE EFFECTS OF PROVIDING KNOWLEDGE OF RESULTS UPON THE PERCEPTUAL DEFENSE EFFECT[1]

MICHAEL J. GOLDSTEIN AND SAMUEL Z. HIMMELFARB

*University of California, Los Angeles*

Recent research on the role of motivational variables in perception has emphasized the need for a closer analysis of the factors operating in a perceptual recognition experiment. The traditional assumption that differences in perceptual accuracy for different classes of stimulus words or pictures reflect the exclusive contribution of the perceptual system has been seriously questioned by a number of writers (Eriksen, 1958; Goldiamond, 1958; Goldstein, 1962). The presence of response biases in articulating the indicator of perception, which exist independent of the presence of a discriminative stimulus, requires that the experimental arrangement for studying perception provide for the statistical or experimental isolation of these response biases from behavior truly under the control of a perceptual stimulus.

A second factor, which has received less attention than the concept of response bias, is the absence of feedback in the perceptual recognition situation. In most experiments, each time the subject makes a guess as to the meaning of an impoverished visual stimulus, he is not provided with any sort of veridical criterion against which to compare his guess. Thus, individual differences in criteria for reporting sensory inputs can be utilized to reinforce certain recognition guesses as being correct and to inhibit other guesses due to their being judged incorrect. This freedom of reinforcement can provide an opportunity for response biases to undergo various sorts of transformations depending upon the subject's reaction to not receiving perceptual feedback. Once again we might misinterpret changes in perceptual accuracy over trials as being due to changes in discriminative ability rather than as systematic changes in response bias due to failure to receive feedback.

Interest in the present problem grew out of an earlier study (Goldstein, 1962) in which provisions were made to observe the perceptual defense effect over trials. Two things were noted in this study: the absolute magnitude of the perceptual defense effect *increased* over trials and the probability of using anxiety words as recognition guesses, which was lower initially than the probability of using neutral words as guesses, *decreased* over trials. The present report is concerned with exploring the degree to which these systematic changes in accuracy and response bias are a result of the lack of feedback from the experimenter during the process of measuring the subject's perceptions.

Specifically, it was predicted that the introduction of knowledge of results early in the series of perceptual responses would lead to: (*a*) a gradual equalization of the probability of using anxiety words and neutral words as recognition guesses and (*b*) that these changes in response bias will be paralleled by a similar equalization over trials of perceptual accuracy scores for anxiety words and neutral words.

## METHOD

*Subjects.* Subjects were 20 UCLA summer session students in a course in psychological aspects of aging given by the senior author. The subjects were older on the average than college sophomores typically used in psychological research, the majority falling between 20–25 years of age. The subjects were run by the junior author and every attempt was made to mask the senior author's interest or participation in the study. Five male and five female subjects were randomly assigned to each experimental group.

[1] The present study was supported by a grant from the UCLA Faculty Research Committee. The paper was worked on by the senior author while a Fulbright research fellow to the Psychological Laboratory, University of Copenhagen, 1960–61.

An abridged version of this paper was presented at the Western Psychological Association meetings, April 1960 at San Jose, California.

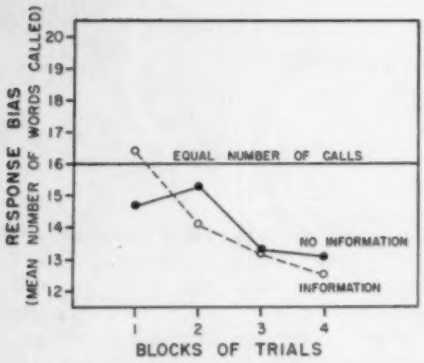



FIG. 1. Mean number of *anxiety* words used as recognition guesses broken down by information group and block of guessing trials.

*Selection of anxiety arousing and neutral words.* The method of selecting four anxiety words and four neutral words for each subject is identical to that reported in an earlier experiment (Goldstein, 1962). Following the selection session, the subject was scheduled to return within 48 hours.

*Perceptual session.* On the second experimental day, two things were done. First, a measure of the subject's visual sensitivity was obtained to determine an exposure time at which the subjects would be approximately 25% correct in his recognition guesses. Second the actual perceptual defense task was run. The method for determining visual sensitivity is also described elsewhere (Goldstein, 1962).

*Perceptual defense measurement.* Following the determination of visual sensitivity, the subject was told that the stimuli to be flashed in the Gerbrands tachistoscope were to be drawn from a set of eight words. The subject was shown the list of eight words, drawn from his word association test, and was given one minute to familiarize himself with them. Following this the subject was instructed:

> This is a study dealing with perception. Here is a list of eight words which will be presented one at a time, in a random order, in this machine. You are to look into the eyepiece and judge which of the eight stimuli is present on each trial by referring to your list. The words will be presented fairly rapidly so that at times it will be fairly difficult for you to see the words and you will frequently have the feeling that you are guessing. Don't be disturbed by this, just guess on every trial.

Following this, the subject was shown the fixation point in the tachistoscope and was instructed to expect the stimuli about this point. After the completion of the instructions, stimulus words, typed on 3 × 5 cards, were exposed one at a time in the tachistoscope according to a prearranged random order for 128 exposures. The only restriction on the random order was that each of the eight words had to appear equally often in each block of 32 trials, a provision which permitted a later analysis of the data over blocks of 32 exposures. All 128 exposure trials were run at the same setting, which was determined in the manner described above.

*Introduction of the knowledge of results factor.* Prior to the perceptual part of the experiment subjects were randomly divided into two groups, an *information group* and a *no information group*. For the first 32 stimulus exposures in the perceptual defense situation neither group of subjects received any information concerning the correctness of their responses. Following the first block of 32 trials, subjects in the Information group were told "Right" every time they guessed correctly. No transition to the information stage was made by the experimenter who merely started providing information on the thirty-third trial with no explanation to the subject. The No Information subjects received no feedback throughout the experiment.

## RESULTS

*Analysis of changes in response bias.* The first hypothesis tested was that the introduction of knowledge of results would result in a gradual equalization of the rate at which anxiety words and neutral words would be used as recognition guesses. To study this, each subject's record was scored for the number of times anxiety words were used as recognition guesses. To observe changes over trials, the total of 128 recognition trials was broken down into four blocks of 32 trials each. Equal

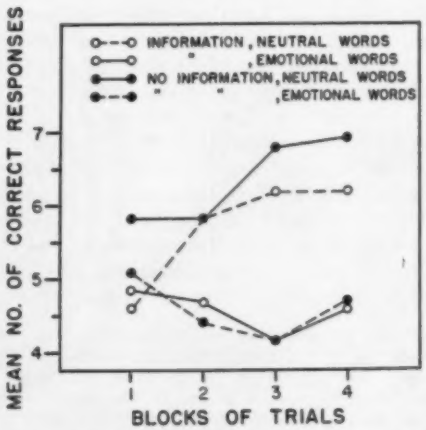



FIG. 2. Mean number of words correctly identified broken down into type of word, information group, and block of guessing trials.

usage of anxiety words and neutral words would be reflected in a score af 16 per block. Figure 1 presents the mean number of times anxiety words were used as recognition guesses in the Information and No Information groups. Inspection of Figure 1 indicates that the probability of using anxiety words as recognition guesses *decreases* in both groups at a very similar rate. The curve for the Information group clearly runs counter to our prediction. The response bias data were analyzed statistically using an analysis of variance of linear trend, in the manner suggested by Grant (1956). This analysis revealed an insignificant $F$ ($F < 1.00$ for $df = 1/18$) for treatment groups which indicates that the linear component of the slope in both curves in Figure 1 do not differ. Knowledge of results has no noticeable influence upon response bias.

*Analysis of accuracy scores.* The second hypothesis was that accuracy scores for anxiety words and neutral words in the Information group would also become more similar over trials. These data are represented in Figure 2. Here the mean number of correct recognitions is presented, broken down by the type of stimulus word, treatment group, and block of trials. Although both treatment groups were treated identically in the first block of 32 trials, Figure 2 suggests an inequality in accuracy on the first block between the two groups. To determine if these first block scores differed significantly from each other, an analysis of variance was done on the first block scores alone. The $F$ between groups (Information vs. No Information) was 1.59 which is not significant at the .05 level for $df = 1/18$. The $F$ for the interaction between treatments and type of stimulus word was 2.02 which also fails to reach statistical significance (for $df = 1/18$) indicating that the four first block scores did not differ significantly among each other. Since this was the case, the analysis of trends over trials used the scores from all four blocks of trials.

Further inspection of Figure 2 also fails to suggest a decrease in the magnitude of the perceptual defense effect over trials. The slopes of the curves for the Information and No Information conditions, although starting from slightly different points of origin, are very similar. An analysis of variance of linear trend was also performed on the accuracy data and the results of the analysis are presented in Table 1. The only significant sources of variance in Table 1 is the type of word (anxiety word vs. neutral word), which means that the slopes of the curves for the neutral words differ significantly from the slopes of the curves for the anxiety words. A further look at Figure 2 reveals that it is the curves for the neutral words in both treatment groups which show a sharp inflection upward, while the anxiety word curves show only a slight downward trend. These data taken in conjunction with the previously reported response bias data suggest that the increase in accuracy for neutral words is accomplished by sharply increasing the rate at which these words are used as recognition guesses, a situation which appears to have little effect upon the accuracy level for anxiety words.

TABLE 1

ANALYSIS OF VARIANCE OF LINEAR SLOPE FOR INFORMATION AND NO INFORMATION GROUPS

| Source | *SS* | *df* | *F* |
|---|---|---|---|
| Between groups: | 909.50 | 20 | |
| Mean | 24.02 | 1 | — |
| Information vs. No Information | 65.02 | 1 | — |
| Error | 820.46 | 18 | |
| Within groups: | 1034.48 | 20 | |
| Type of word | 308.02 | 1 | 8.101* |
| Word × Treatment | 42.02 | 1 | |
| Error | 684.44 | 18 | |

* $p < .05$.

## DISCUSSION

The main hypothesis of the present study is clearly not supported; knowledge of results does not modify accuracy or response bias in a perceptual defense experiment. Evidence still exists from the progressive shifts in accuracy and response bias for a self-reinforcing process which modifies accuracy and probability of response, but the locus of the reinforcement is not clear. It is obviously not a simple case of the subject not having a veridical criterion against which to compare his recognition guesses.

Some clarification may come from attempting to unravel the direction of the causal re-

lationship between the changes in accuracy and the changes in response bias noted over trials. Two possibilities exist: The accuracy changes are primary and the changes in the response bias are secondary effects of the accuracy changes. The changes in response bias are primary and the changes in accuracy are an artifact of the changes in probability of usage of the two classes of words as recognition guesses. Evidence from an earlier study (Goldstein, Himmelfarb, & Feder, 1962) tend to support the second hypothesis. In this study, the perceptual recognition experiment was changed so that the subject indicated his perception by locating the stimulus in space, rather than by articulating the stimulus word, as was the case in the present study. The use of the spatial forced-choice indicator prevents response bias from affecting perceptual recognition scores. In the forced-choice study, no changes in accuracy over trials were noted for either class of stimulus word. This finding suggests that the changes in accuracy noted in the present study are an artifact of systematic shifts in response bias. The subject is restricting his recognition guesses progressively to that subset of responses termed neutral words, and is, in this fashion, increasing the probability of a chance congruence in accuracy for the neutral words.

If the hypothesis that it is response bias and not accuracy which shifts over trials, should be supported by other research, then how can we account for this type of shift? Two hypotheses are tenable, one which emphasizes the presence of a perceptual stimulus and a second which emphasizes response variables exclusively. The first hypothesis centers about the type of implicit verbal reinforcement which the subject is believed to administer after using each class of words as recognition guesses. It has been shown (Goldstein, 1962), that anxiety words have a lower probability of evocation than neutral words, when a blank stimulus is flashed in a tachistoscope. It is conceivable that the subject feels less subjective confidence after using a low probability word as a response than when he uses a high probability word. These differences in subjective confidence could affect the rate at which the subject uses these two classes of words as responses over a series of trials. The second possibility is that this behavior is not under control of the stimulus at all. Let us forget for a moment that a real perceptual stimulus was presented to the subject and assume that the task assigned to him is to label some objects. He is given two classes of labels and told that he is to use them to label different objects. One class of labels has been associated with anxiety in the past while the other has not. What we then have is a classic avoidance learning paradigm. The subject has the choice of emitting an anxiety linked word or a neutral word on each trial. The anxiety linked word is painful for him to emit, so therefore he emits the neutral word whenever possible. Each time he uses a neutral word as a response, he is avoiding using an anxiety word, a situation closely paralleling many avoidance learning experiments. If we assume that not using an anxiety linked word is reinforcing of his already existing avoidance tendencies for that word, then each time the subject uses a neutral word, he is increasing the likelihood of using that word again because it is an anxiety reducing avoidance response. Thus, the progressive increase in the use of neutral words and decline in the use of anxiety words is explainable from an avoidance conditioning paradigm and does not even require the presence of a visual stimulus. Experiments, currently underway, using blank stimulus groups are directed at choosing between these two hypotheses. Should the same sort of progressive changes in response bias show up with no true discriminative stimulus present, then Hypothesis 2, which emphasizes reinforcement of intraverbal habits, would be reasonably well supported.

Regardless of which hypothesis is supported, it should be apparent from the present research that the perceptual recognition experiment does not merely reveal stable reaction tendencies, but is itself a dynamic situation in which sequences of responses are modifying pre-existing reaction tendencies.

## SUMMARY

The present paper reported a study designed to observe the perceptual defense

effect over trials. It was hypothesized that the introduction of knowledge of results would result in an equalization of the probability of using anxiety words as responses and in the accuracy for neutral and anxiety words. Both hypotheses failed to be confirmed. The probability of using anxiety words as recognition guesses *decreased* over trials and the accuracy curves diverged over trials.

Discussion centers about specifying the factors responsible for the systematic changes observed over trials.

## REFERENCES

ERIKSEN, C. W. Unconscious processes. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1958*. Lincoln: Univer. Nebraska Press, 1958. Pp. 169–227.

GOLDIAMOND, I. Indicators of perception: I. Subliminal perception, subception, unconscious perception: An analysis in terms of psychophysical indicator methodology. *Psychol. Bull.*, 1958, **55**, 373–411.

GOLDSTEIN, M. J. A test of the response probability theory of perceptual defense. *J. exp. Psychol.*, 1962, **63**, 23–28.

GOLDSTEIN, M. J., HIMMELFARB, S., & FEDER, WALDA. A further study of the relationship between response bias and perceptual defense. *J. abnorm. soc. Psychol.*, 1962, **64**, 55–62.

GRANT, D. A. Analysis of variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

(Received February 13, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 148–150

## *CRITIQUE AND NOTES*

# SEATING ARRANGEMENT AND LEADERSHIP EMERGENCE

LLOYD T. HOWELLS AND SELWYN W. BECKER

*Graduate School of Business, University of Chicago*

In task oriented discussion groups a leader will tend to emerge as a result of interaction among the group members. Bavelas (1950) and Goldberg (1955) have concluded that the size of the group and the type of communication pattern used are determiners of leadership emergence. Heise and Miller (1951) presented evidence to support the hypothesis that the most centrally located member of a group is likely to emerge as leader. Bales (1953) also found after analyzing communication patterns and leadership that "the communication pattern tends to 'centralize' . . . around a leader through whom most of the communication flows." Excluding Bales' analysis the experimental conditions in these studies were characterized by controlled communication patterns where the number of channels and the direction of communication were fixed.

With free communication possible, i.e., where all channels are available to all members of the group, Festinger, Schacter, and Back (1950) found that communication among individuals tends to be maximized between individuals positioned opposite to one another and conversely, that there is less communication between people placed side by side. A similar conclusion was drawn by Steinzor (1950) who stated that "the greater . . . the expressive stimulus value a member of a group has for others the more nearly opposite he sits from one in a circle" (p. 554). Because his subjects were seated in a circle, it was also possible for Steinzor to conclude that a participant who had a greater mean seating distance from all the other participants would be likely to attain higher leadership status.

Bass and Klubeck (1952) partially tested Steinzor's (1950) conclusions by arranging subjects on each side of a V shaped arrangement. With this arrangement the mean seating distances varied and following Steinzor it was expected that the subjects on the open end of the V would more frequently emerge as leaders. The data did not conform to this expectation and Bass concluded that the "particular seat a person occupied was of negligible importance in determining a participant's tendencies to attain leadership status . . ." (p. 727).

From the Festinger et al. (1955) and Steinzor (1950) results one could deduce that the communications of a small group sitting at a table will tend to flow "across" rather than "around" the table. Further, from the Bevelas' (1950) and Bales' (1953) results it can be stated that the individual who controls communication, i.e., the individual who sends and receives more messages, will emerge as leader.

If these two conjectures are correct then, despite Bass' (1953) failure to find the effect, seating arrangements can affect communication flow which, in turn, will influence leadership emergence. Accordingly, the general hypothesis to be tested in the present study is that *seating position, by influencing the flow of communication, is a determiner of leadership emergence.*

## METHOD

*Subjects.* One hundred undergraduate men enrolled at the University of Chicago volunteered in order to earn $1.00 as experimental subjects. Twenty groups of five were constituted by random selection from the pool of 100 subjects with the restriction that all members of a group be unable to recognize the names of the remaining group members.

For purposes of identification as each group member appeared for his evening session he was given a colored wrist band (Blue, Red, Green, White, or Yellow) to which he had been randomly assigned and instructed to sit in the chair with the matching color. The colors on the chairs were rotated so that over the 20 sessions each color was assigned to each chair position an equal number of times. The chairs were arranged so that two were evenly spaced along one of the long sides of a $6' \times 3'$ table and three were evenly spaced on the opposite, long side of the table. The subjects were informed that the purpose of the experiment was to investigate the effects of various communication patterns on creativity after which the instructions pertinent to the task were given.

The instructions included an explanation of the task, anagrams, and the rules for evaluating the group's production. The subjects were told that each word was worth a number of points and that some words were worth more than others, and furthermore that as they generated words it would be possible to establish a hypothesis about the kinds of words which have greater value. They were

also told that a clock was set at random so that the alarm signaling the end of the session would ring sometime between 15 and 45 minutes after they started working. In reality all groups worked for 25 minutes. At the alarm the group was required to submit the product of their effort on one sheet of paper.

The task selected, and the conditions imposed on its completion, were designed to stimulate group interaction and to enhance the chances for development of group structure. That is, it was possible for one member to designate another as recorder so the requirement of submitting "one sheet representing the group's effort" would be satisfied. Decisions had to be made on the division of time and labor with regard to generating words and looking for the clue.

When the alarm rang ending the work session, questionnaires were distrbuted and completed by the group members. The questionnaires required each subject to rank the other four subjects, but *not* himself, on 15 personality characteristics and also to compare each subject with all the other subjects, again excluding himself, for leadership qualities demonstrated during the session. Since it has been shown that nominations by group members are more reliable leadership measures than are observer ratings or other leadership criteria (Carter, Haythorn, & Howell, 1950), the subjects' leadership ratings were taken as the criteria of leadership. Where there was a 4–1, or 3–2 vote favoring a single person, he was considered the leader of the group. Any other vote was deemed inconclusive and the leader was assumed to be on the side of the table mitigating against the hypothesis.

The procedures provided for two subjects to face three subjects across a table and engage in a problem solving task. Based on the theoretical considerations described above it was expected that communication would flow across rather than around the table. Thus, it was expected that each of two subjects would exert influence on three subjects and each of three subjects would exert influence on two subjects. Leadership ratings made by the members of the group provided data for a test of the following hypothesis:

*A greater number of leaders than would be expected by chance will emerge from the two-seat side of the table.*

## RESULTS AND DISCUSSION

The analysis of the data was based on the assumption that each person in the group had an equal chance of being voted leader. As such, it would be expected that 8 leaders would emerge from the two-seat side of the table and 12 from the three-seat side. Since the observed results are in the predicted direction and differ significantly from those expected by chance it is concluded that seating position influences communication and so is a determinant of leadership emergence in a task oriented group.

TABLE 1

NUMBER OF LEADERS EMERGING FROM TWO- AND THREE-SEAT SIDES OF TABLE

| Two-seat side | Three-seat side |
|---|---|
| 14 | 6 |

Note.—$\chi^2 = 7.5$, $p < .01$.

Given these data it seems possible to reconcile the differences between Steinzor's (1950), Bass and Klubeck's (1952), and the present study. Steinzor found that a person sitting opposite another around a circle provided more stimulus for that other and also that the largest mean seating distance was a determiner of leadership. That the second conclusion may be viewed as an artifact of the circular arrangement of subjects seems possible, especially since Bass and Klubeck failed to support the conclusion with subjects arranged around a V shaped arrangement. If "oppositeness" rather than seating distance is important then Bass and Klubeck could have expected no differences in their situation, since no subject had an advantage in the number of other subjects he faced. Differences could be expected in the situation reported here, i.e., two subjects opposite three subjects, for each of two could possibly influence each of three.

## SUMMARY

Based on the proposition that seating position effects the flow of communication which, in turn, effects leadership emergence 20 groups of five subjects performed a problem solving task and then rated the group members on exhibited leadership. The seats were arranged so that two subjects were opposite three subjects. A greater number of leaders than would be expected by chance emerged from the two-seat side of the table, thus, lending support to the hypothesis.

## REFERENCES

BALES, R. F. The equilibrium problem in small groups. In T. Parsons, R. F. Bales, & E. A. Shils (Eds.). *Working papers in the theory of action*. Glencoe, Ill.: Free Press, 1953. Pp. 111–161.

BASS, B. M., & KLUBECK, S. Effects of seating on leaderless groups. *J. abnorm. soc. Psychol.*, 1952, **47**, 724–727.

BAVELAS, A. Communication processes in task-oriented groups. *J. Acoust. Soc. Amer.*, 1950, **22**, 725–730.

CARTER, L. F., HAYTHORN, W., & HOWELL, M. A further investigation of criteria of leadership. *J. abnorm. soc. Psychol.*, 1950, **45**, 350–360.

FESTINGER, L., SCHACHTER, S., & BACK, K. *Social processes in informal groups: A study of a housing project*. New York: Harper, 1950.

GOLDBERG, S. C. Influence and leadership as a function of groups structure. *J. abnorm. soc. Psychol.*, 1955, **51**, 119–122.

HEISE, G., & MILLER, G. Problem solving by small groups using various communication nets. *J. abnorm. soc. Psychol.*, 1951, **46**, 327–335.

STEINZOR, B. The spatial factor in face to face discussion groups. *J. abnorm. soc. Psychol.*, 1950, **45**, 552–555.

(Received January 17, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 151–154

# THE EFFECTS OF PATIENTS' RELATIONSHIPS WITH PEERS AND STAFF ON THEIR PYSCHIATRIC TREATMENT PROGRAMS [1]

E. ROBERT SINNETT
*University of Kansas*
AND DANA B. HANFORD
*Northern State Hospital, Sedro-Woolley, Washington*

Psychiatric treatment of the individual patient is influenced by social-psychological factors other than his need for therapy and the type and severity of his condition as denoted by psychiatric judgments of diagnosis and prognosis. Hollingshead and Redlich (1958), Hunt (1959), and others have provided convincing evidence that membership in a given social class and the characteristic psychological orientation associated with that class are substantially related to the treatment of the psychiatric patient.

The present study was an exploratory probe of the relationship between the affectional responses elicited by a patient from his ward physician and his peers and the type of treatment program he received. One objective was to investigate whether a physician's degree of like or dislike for a patient influenced the patient's assignment to individual psychotherapy.[2] Similarly, by means of sociometric data, the ties a patient had to his peers were examined for their possible relationship to his receiving individual psychotherapy. For purposes of comparison, however, both socioeconomic indices and psychiatric judgments of diagnosis and prognosis were included. Thus, three sets of variables were analyzed in relation to a patient's assignment to individual psychotherapy: a patient's relationships to his fellow patients and to his ward physician; socioeconomic indices (educational and occupational levels); and psychiatric judgments of diagnosis and prognosis.

## METHOD

*Subjects.* The subjects were 85 male veteran psychiatric patients on an open ward intensive treatment service. This service consisted of four 24-bed wards which were parallel, quite close to each other, and joined by a common corridor. Social contacts among patients were not limited to members of a single ward since most activities both planned and informal were common for all four wards.

The diagnostic composition of the group was: psychoses, 48%; neuroses, 36%; personality disorders, 10%; and neurological disorders, 6%. The age range was from 19 to 67 with a mean of 37.5 years. The mean rating of educational level (Warner, Meeker, & Eells, 1949) was slightly lower than completion of high school. The mean rating of occupational level (Warner et al., 1949) was at the level of skilled workers.

The median length of current, continuous hospitalization for the subjects at the time of the study was 8 months, and the range was from 1 to 123 months. In order to assure each patient a reasonable time unit in which relationships might be established, a minimum of 1 month current hospitalization was used as a criterion for inclusion in the study. Subsequent analysis of the data showed that sociometric status was unrelated to length of hospitalization ($\chi^2 = 3.97$, $df = 3$, *ns*).

*Sociometric questionnaire.* The sociometric questionnaire constructed for this study utilized free choice. That is, in responding to eight questions such as, "With whom would you like to be in group therapy?" and "With whom do you prefer not to associate?" a subject was free to nominate as many or as few fellow patients as he chose. Choices were not restricted to a particular ward or service.

Ninety-three percent of the subjects returned their questionnaires. Twenty-one percent of the total group returned questionnaires signed but without responses. Failure to submit a questionnaire or returning a blank questionnaire was originally considered a basis for excluding individuals from the study. However, examination of the data showed that both these groups received statistically fewer choices from their peers ($\chi^2 = 17.01$, $df = 2$, $p < .001$). Therefore, it was decided to regard both nonresponding classes as expressing meaningful social behavior in their nonresponse rather than to bias the sample by excluding these less popular subjects. The number of rejections received from peers did not differentiate these classes of nonrespondents from those completing the questionnaire ($\chi^2 = .62$, $df = 2$, *ns*).

*Interview and rating procedures.* The patients' ward physicians, who were second or third year

[1] Research was done at Veterans Administration Hospital, Topeka, Kansas.

[2] Patients did not receive individual psychotherapy from their own ward physician. Hence this study is not concerned with the liking of the therapist for his patients but with the affective responses accorded the patient in his hospital environment by peers and ward physician.

TABLE 1

RELATIONSHIP BETWEEN SOCIOMETRIC CATEGORY AND ASSIGNMENT TO INDIVIDUAL PSYCHOTHERAPY

| Sociometric category | Treatment | | |
|---|---|---|---|
| | No individual therapy | Individual therapy | *N* |
| Popularity | | | |
| High | 9 | 16 | 25 |
| Average | 23 | 9 | 32 |
| Rejects | 6 | 2 | 8 |
| Isolates | 14 | 6 | 20 |
| Totals | 52 | 33 | 85 |

Note.—$\chi^2 = 9.50$, $df = 3$, $p < .05$.

psychiatric residents, were interviewed for information about the patients' treatment programs and their progress in treatment. The four psychologists who did the structured interviews with the physicians felt that a condition of good rapport existed since most of them had regular, sustained contact with the residents as part of their clinical duties. At the time of the interviews the residents had known their patients for from 1 to 6 months. These 13 physicians were also asked to rate their degree of liking or disliking for each of their patients on a seven-point scale.

Diagnosis from the final summary and prognosis from the initial report of examination (based on a 2-week evaluation) as well as information about length of hospitalization and milieu treatment program were obtained from the patients' records.

## RESULTS

The statistical analyses used chi square throughout. Three approaches were used to analyze the sociometric data. First, individuals were categorized as relatively isolated; relatively rejected; popular, or of medium popularity on the basis of the pattern of choices made and received; rejections made and received; and the number of mutual or reciprocal choices (i.e., persons who chose one another). Distributions of each of these five variables were dichotomized or trichotomized (where feasible) in order to classify subjects. For example, the isolates, as defined in this study, were those who made and received relatively few choices or rejections and had few mutual choices. The rejected individuals were those who received a relatively high number of rejections. It must be kept in mind that these categorizations are not absolute. For example, it happened that no patient who received a high number of rejections received *no* choices. That is, the most rejected patients were not unpopular but rather, controversial. Therefore the data were analyzed not only by sociometric categories, but also, as a second method of analysis, by the component scores (number of choices made, rejections received, etc.) from which the categories were derived.

The set of sociometric categories was found to be significantly related to whether or not a patient was given or had been given individual psychotherapy (see Table 1). Relatively popular patients were more likely to receive individual psychotherapy than less popular patients.

Analysis of the sociometric questionnaire by component scores revealed that patients who were treated by individual psychotherapy received more choices from their peers ($\chi^2 = 3.84$, $df = 1$, $p < .05$); they also made more choices ($\chi^2 = 5.12$, $df = 1$, $p < .05$) and made *more* rejections ($\chi^2 = 7.61$, $df = 1$, $p < .01$). They also received *more* rejections from their peers but not to a significant degree ($\chi^2 = 2.39$, $df = 1$, $p < .20$). As an outcome of these analyses a factor of affect intensity of peer relationships appeared to be related to receiving individual psychotherapy. In order to construct a combined measure of intensity of ties to peers as a third approach to data analysis, number of choices made, number of choices received, number of rejections made, and number of rejections received were dichotomizd at their medians. Scores for each of these variables were treated as follows: values above the median were given a weight of 1; those below the median, 0; and these weights were summated for each subject over the four variables. The resultant affect intensity score was found to be positively related to receiving individual psychotherapy ($\chi^2 = 5.27$, $df = 1$, $p < .05$). Thus, both expressing and eliciting positive or negative feelings to or from one's peers were associated with receiving individual psychotherapy.

Physician liking was positively related to receiving individual psychotherapy; i.e., patients liked by their ward physician were more likely to receive this form of treatment (see Table 2).

TABLE 2

RELATIONSHIP BETWEEN BEING LIKED BY ONE'S PHYSICIAN AND ASSIGNMENT TO INDIVIDUAL PSYCHOTHERAPY

| Physician liking | Treatment | | |
|---|---|---|---|
| | No individual therapy | Individual therapy | *N* |
| Liked by physician | 34 | 31 | 65 |
| Not liked by physician | 18 | 2 | 20 |
| Totals | 52 | 33 | 85 |

Note.—$\chi^2 = 9.13$, $df = 1$, $p < .01$.

It should be noted that the practice in the hospital setting where the study was conducted was for ward physicians *not* to take their own patients in psychotherapy. Assignment to individual psychotherapy was characteristically a product of a joint decision of a ward physician and his supervisor, or an outcome of a staff conference where the resident presented his psychiatric case study of a particular patient.

Although the relationship between physician liking and receiving psychotherapy is open to the interpretation that therapy made patients more likeable rather than being a factor in selection for treatment, there is some evidence that would tend to be incongruent with the former interpretation—namely, the fact that among the 33 subjects who received psychotherapy, physician liking showed a low, *negative*, insignificant relationship ($r = -.16$) with number of months in psychotherapy.

An additional finding from this study supports the interpretation that patients who are ignored by or isolated from their peers by their own choice receive a less intensive treatment program. Physician's prescriptions for assignments of patients to activities tended to include fewer stated treatment objectives for isolates and also for rejects than for patients of higher popularity (the Kruskal-Wallis $H$ test yielded a $\chi^2$ value of 6.38, $df = 3$, $p < .10$).

Consistent with findings of a previous study using subjects drawn from the same population (Sinnett & Hammersley, 1957), prognosis on admission was unrelated to receiving individual psychotherapy ($\chi^2 = 0.11$, $df = 1$, *ns*). In the earlier study diagnosis was also unrelated to receiving individual psychotherapy. However, in contrast with the earlier finding, the present study showed that diagnosis and receiving individual psychotherapy were related: neurotics were more likely to receive this form of treatment ($\chi^2 = 10.27$, $df = 3$, $p < .02$).

Although the population receiving individual psychotherapy had a higher educational level, ($\chi^2 = 5.56$, $df = 1$, $p < .02$), occupational level was not so related in this sample ($\chi^2 = 0.32$, $df = 1$, *ns*). The findings from the previous sample (Sinnett & Hammersley, 1957) were that both the more highly educated and those who had held higher level jobs more often received individual psychotherapy.

Although one might anticipate that degree of isolation, rejection, and popularity would be related to diagnosis and prognosis—for example, that schizophrenics might predominate among the isolates—no such relationship was found (see Table 3). Similarly, diagnosis and prognosis were not related to being liked by one's physician ($\chi^2 = 1.05$ and 0.00, $df = 3$ and 1, respectively) even though in some cases the three judgments of liking, diagnosis, and prognosis might have been made by the same physician. Thus, it appears that no generalized halo effect confounded these judgments.

TABLE 3

RELATIONSHIP BETWEEN SOCIOMETRIC CATEGORY AND DIAGNOSIS

| Sociometric category | Diagnosis[a] | | | | |
|---|---|---|---|---|---|
| | Brain damaged | Psychotics | Personality disorders | Neurotics | *N* |
| Popularity | | | | | |
| High | 1 | 10 | 4 | 10 | 25 |
| Average | 1 | 16 | 4 | 11 | 32 |
| Rejects | 2 | 5 | 0 | 1 | 8 |
| Isolates | 1 | 9 | 1 | 9 | 20 |
| Totals | 5 | 40 | 9 | 31 | 85 |

Note.—$\chi^2 = 3.42$, $df = 3$, *ns*.

[a] Because of many small cell frequencies, the categories were combined as follows for the $\chi^2$ test: brain damaged subjects were excluded and neurotics and personality disorders were combined.

Somewhat in contrast to what has been found by Hollingshead and Redlich (1958), indices of socioeconomic level did not show a substantial relationship to physician liking. Being liked by one's physician was not related to occupational level ($\chi^2 = 1.27$, $df = 1$, *ns*). However, there was a trend toward a positive relationship between being liked by one's physician and educational level ($\chi^2 = 3.19$, $df = 1$, $p < .10$). The more highly educated patients were also more popular with their peers ($\chi^2 = 8.59$, $df = 3$, $p < .05$). There was a trend toward a relationship between sociometric category and physician's regard for a patient: patients who are relatively popular with their peers tend to be liked by their physicians ($\chi^2 = 6.29$, $df = 3$, $p < .10$). The relationship between physician liking and status among patient peers was explored further using the affect intensity measure described above. Expressing and eliciting positive and/or negative feelings to or from one's peers was positively related to being liked by one's physician ($\chi^2 = 3.85$, $df = 1$, $p < .05$). Conversely, those who made fewer choices and rejections and were infrequently chosen or rejected by peers were less liked by their physicians, just as they were less often selected for psychotherapy.

The sociometric categories were independent of age ($\chi^2 = 2.07$, $df = 3$, *ns*) and length of hospitalization ($\chi^2 = 3.97$, $df = 3$, *ns*). Physician liking also was not related to length of hospitalization ($\chi^2 = 0.03$, $df = 3$, *ns*).

## DISCUSSION

It appears that a relatively intense affective responsiveness leads to a patient being considered a good candidate for psychotherapy. The degree to which affective responsiveness is an attribute necessary to the conduct of psychotherapy and the degree to which there are simply personal preferences among psychiatrists for working with more responsive patients are matters that are open to question.

To date the relationships of sociometric measures and physician liking to receiving individual psychotherapy have not been cross-validated. However, the findings of Hollingshead and Redlich (1958, pp. 344–345), imply a positive association between liking and selection for psychotherapy. They show that physicians like better those patients from higher social classes and that patients from higher social classes are more likely to receive psychotherapy.

Reciprocal or mutual choice behavior among the subjects of this study was strikingly similar to the results obtained by Murray and Cohen (1959) in their open ward psychiatric patient group, even though somewhat different sociometric measures were utilized in the two experiments. They found 50% of their patients had one or more reciprocal positive choices and the corresponding figure for the present study was 39%. One or more mutual negative choices or rejections were found by Murray and Cohen in 6% of their cases as compared with 0% in this study.

Although previous studies (Davids & Parenti, 1958; McMillan & Silverberg, 1955; Murray & Cohen, 1959; Northway & Widgor, 1947) and a review (Lindzey & Borgatta, 1954) have shown less consistency than might be desired, generally subjects receiving relatively few choices or relatively many rejections have been judged on the basis of other measures to be more maladjusted. In this study, however, there was no association found between diagnosis or prognosis and sociometric measures.

Ordinarily in the appraisal of patients a physician's feelings toward the patient and the patient's relationships with his peers are subordinated to "objective" considerations of diagnosis and prognosis; nevertheless, the matrix of interpersonal relations appears to make itself felt in the patient's treatment program.

The psychiatric patient comes to the hospital for treatment rather than be treated in his home environment mainly in order that his social climate may be controlled so that he can achieve a perspective view of problems with himself and in his relationships with others. Yet, should he manage to duplicate among his patient peers and professional staff the climate he has left at home, the effectiveness of efforts to modify his pattern of relationships with others may be less than is supposed.

The findings appear to lend support to the current argument that the whole nexus of patient relationships—the social environment of the mental hospital—ought to be examined and treated as well as the problems of individual patients.

## SUMMARY

The major finding of this research is that a patient's relationships with his peers and his physician relate to the type of treatment program he receives as well or better than indices of his social class or his psychiatric diagnosis or prognosis. Among a group of 85 male veteran patients, the subjects most often selected for psychotherapy were: patients who were most popular among their peers; patients whose affective ties (both positive and negative) with peers were strongest; and patients most liked by their ward physicians. Conversely, the patient who was neither "in with" nor a controversial figure among his fellow patients, who was either only average in popularity or isolated from his peers, was not often selected for psychotherapy.

## REFERENCES

DAVIDS, A., & PARENTI, ANITA N. Personality, social choice, and adults' perceptions of these factors in groups of disturbed and normal children. *Sociometry*, 1958, **21**, 212–224.

HOLLINGSHEAD, A. B., & REDLICH, F. C. *Social class and mental illness*. New York: Wiley, 1958.

HUNT, R. G. Socio-cultural factors in mental disorder. *Behav. Sci.*, 1959, **4**, 96–106.

LINDZEY, G., & BORGATTA, E. Sociometric measurement. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. 1. *Theory and method*. Cambridge: Addison-Wesley, 1954. Pp. 405–448.

MCMILLAN, J. J., & SILVERBERG, J. Sociometric choice patterns in hospitalized ward groups with varying degrees of interpersonal disturbances. *J. abnorm. soc. Psychol.*, 1955, **50**, 168–172.

MURRAY, E. J., & COHEN, M. Mental illness, milieu therapy, and social organization in ward groups. *J. abnorm. soc. Psychol.*, 1959, **58**, 48–54.

NORTHWAY, MARY L., & WIDGOR, BLOSSOM T. Rorschach patterns related to the sociometric status of school children. *Sociometry*, 1947, **10**, 186–199.

SINNETT, E. R., & HAMMERSLEY, D. W. The relationship between attributes of patients selected for psychotherapy and improvement measures. *Amer. Psychologist*, 1957, **12**, 432. (Abstract)

WARNER, W. L., MEEKER, MARCHIA, & EELLS, K. *Social class in America*. Chicago: Science Research, 1949.

(Received January 12, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 155–159

# BIRTH ORDER, NEED ACHIEVEMENT, AND CONFORMITY

EDWARD E. SAMPSON

*University of California, Berkeley*

Although Alfred Adler (1945) suggested the importance of an individual's ordinal position in the family, until recently, very little research had been conducted using ordinal position as an independent variable. A few years ago, Schachter (1959) published a book in which he reported the interesting results of a series of studies on ordinal position, motivation, and behavior in social or experimental situations. In his work, Schachter suggested the hypothesis that first born individuals were more anxious and affiliatively dependent than later born persons. This suggests that the first born individual should be more prone to influence by others; and, in fact, Schachter reports the findings of a study by Ehrlich that supports this hypothesis. Ehrlich found that first born males were less resistant in a social influence situation than later born males.

The data reported in this paper derive from a series of three experiments, two of them directly concerned with social influence, the third more directly concerned with the need for achievement (n Ach).[1] As a routine part of each of these experiments, birth order data were obtained from each subject and analyzed in order to test the relationships between birth order and conformity in two experiments and birth order and n Ach in one experiment.

## Method

### *Birth Order*

Information was obtained about the individual's order of birth by a simple question on which each subject indicated whether he was the first born, second born, etc. The data were classified into two major groupings: first born versus later born. The former category consisted of first and only and first with sibs. The latter category consisted of all persons second born or later. For the analyses of this study, such factors as age distance between self and sibs or sex of nearest sibs were not controlled because of the small *N*.

### *Experimental Situations*

There were two experimental situations which yielded data about conformity and birth order. The first situation, A, consisted of an audience-type influence experiment in which a hired assistant, playing the role of a debator, sought to influence the subject to adopt the position which he advocated on an attitude issue about Russia and the United States. The experiment used a $2 \times 3$ design, with a condition of negative attraction towards the debator and one of neutral attraction. Within each of these, the debator, was a false-expert—that is, he based his influence on an assumed expertness which he actually did not possess; sought to control the subjects' behavior—that is, he told the subjects that he possessed the ability to completely determine their beliefs in this particular situation; or was neither—that is, the control groups. The measures of conformity consisted of the direction and distance of the subjects' change from their preinfluence to their postinfluence position.

The second experimental situation, B, consisted of a four-man group working in separate cubicles on a simple performance task. The task involved drawing circles with dots in the middle for a series of 10-second trials. During the course of the experiment, an individual assumed to be the leader of this four-man group gave instructions to the subjects to increase their performance. The leader promised rewards for conformity to his instructions to half the groups and threatened fines for nonconformity to the other half. The rewards and fines consisted of points given or taken away, the points being worth money. The group leader also gave each subject his evaluation of their performance on the task. Upon receiving his group leader's evaluation, each subject was asked to record his own personal evaluation on the same scale employed by the group leader. The measure of conformity in this situation consisted of the discrepancy between the subject's own personal evaluation and the group leader's evaluation of his performance. Thus, as with the conformity measure in Situation A, this involved a judgmental or cognitive conformity rather than an action conformity.

The third experimental situation, C, gave information about birth order and n Ach. It consisted of a digit-symbol task given to groups of persons premeasured on n Ach and test anxiety. The major concern in this situation was to determine differences in digit-symbol performance between persons motivated by the need to achieve as compared with persons motivated to avoid failure.

[1] The author wishes to thank John R. P. French, Jr. for his invaluable contributions in the first experiment, and Ernest Harburg for his collaboration in the third experiment, which he is now preparing for eventual publication.

TABLE 1

THE RELATION BETWEEN BIRTH ORDER AND THE NEED FOR ACHIEVEMENT

| Total | Birth order | | Total | Birth order | |
|---|---|---|---|---|---|
| | First | Later | | First | Later |
| High n Ach | 23 | 7 | High Ach-Low TAQ | 13 | 2 |
| Low n Ach | 15 | 16 | Low Ach-High TAQ | 6 | 9 |
| | $\chi^2 = 4.05^a$, $p < .05$ | | | $p = .05$ by Fisher's exact test | |
| Males | | | | | |
| High n Ach | 12 | 3 | High Ach-Low TAQ | 6 | 0 |
| Low n Ach | 9 | 7 | Low Ach-High TAQ | 3 | 3 |
| | $\chi^2 = 1.03^a$ | | | | |
| Females | | | | | |
| High n Ach | 11 | 4 | High Ach-Low TAQ | 7 | 2 |
| Low n Ach | 6 | 9 | Low Ach-High TAQ | 3 | 6 |
| | $\chi^2 = 2.17^a$ | | | | |

[a] Corrected for continuity.

### *Subjects*

In Experimental Situation A, involving an attitude issue about Russia and the United States, there were 88 female subjects obtained from introductory psychology courses at the University of Michigan. In Experimental Situation B, involving four-man groups, there was a total of 116 male subjects obtained from recruits at the Coast Guard base in Oakland, California. The final set of subjects in Experimental Situation C consisted of 31 males and 30 females who were members of an undergraduate social psychology class at the University of Michigan.

### *Measurement of n Ach*

n Ach was measured by the projective French (1958) Test of Insight.[2] Test anxiety was measured by the Mandler-Sarason test anxiety items (Mandler & Cowen, 1958). These measures were obtained only on persons in Experimental Situation C. The inclusion of the measure in test anxiety enables a further refinement to be made of n Ach and fear of failure. Subjects above the median on n Ach and below the median on test anxiety constitute one "pure" group in contrast to the subjects below the median on n Ach and above it on test anxiety. This is the type of breakdown suggested by Atkinson (1958; Atkinson & Litwin, 1960).

## RESULTS [3]

### *Birth Order and n Ach*

The analyses presented in Table 1 permit us to look at the achievement measure alone and the split involving both achievement and anxiety (Experimental Situation C). For the total sample of both men and women, there is a significant tendency for first born persons to have a higher need for achievement than later born persons. This relation is supported in the finer breakdown, employing both the measure of achievement and the measure of anxiety. Both of these relationships are significant at the .05 level. The trend of the relationship is the same for both males and females taken as separate groups, but is not statistically significant ($\chi^2 = 1.03$ for the males; $\chi^2 = 2.17$ for the females). There is, however, a slight, nonsignificant indication that this relationship is stronger for females than for males.

### *Birth Order and Conformity*

Experimental Situations A and B each involved an influence situation in which we could measure the conformity or resistance of first and later born persons. Table 2 presents the relevant data from Experimental Situation A. The statistical test is significant at the .10 level, with the direction of the relationship suggesting that first born persons have higher resistance to influence than later born persons.[4] This finding is directly opposite to that Ehrlich has reported by Schachter (1959), but his subjects were males.

Although the sample on which these conformity data were obtained is not the same as that on which the n Ach data were obtained, the groups being compared in each are female undergraduate college students at the University of Michigan.

[2] Ernest Harburg coded the projective test data measuring n Ach.

[3] All statistical tests are two-tailed, as we have no firm basis for making specific directional predictions.

[4] Because Experimental Situation A employed an experimental division of groups into those having a neutral debator and those having a disliked debator, we conducted an analysis of birth order and conformity within these two major groupings. For both those having a neutral debator and those having a disliked debator, there is an indication (not statistically significant, however) that later born persons are less resistant than first born persons. This finding is consistent with the findings reported in Table 2 for the total sample.

TABLE 2

THE RELATION BETWEEN BIRTH ORDER OF FEMALES AND RESISTANCE TO INFLUENCE ON AN ATTITUDE ISSUE

| | Birth order | |
|---|---|---|
| | First | Later |
| Low resistance to influence | 18 | 29 |
| High resistance to influence | 24 | 17 |
| | $\chi^2 = 3.60$, $p < .10$ | |

TABLE 3

THE RELATION BETWEEN BIRTH ORDER OF MALES AND RESISTANCE TO ACCEPTANCE OF THE GROUP LEADER'S EVALUATION[a]

| | Birth order | | t |
|---|---|---|---|
| | First | Later | |
| Reward conditions | 3.62 | 5.67 | 1.86[b] |
| Fine conditions | 5.07 | 5.88 | .88 |

[a] The measure reported is the mean discrepancy between the group leader's evaluation of the subject's performance and the subject's own evaluation of his performance.
[b] Significant at .06 level.

and are, thus, a relatively homogeneous sample. If we look at these results on conformity and the results on n Ach, we note that they are consistent, but only the former approaches significance. That is, in general, first born females are more resistant in a social influence situation and tend to be higher in the need for achievement than later born females.

Table 3 presents the data from Experimental Situation B on birth order and conformity for the Coast Guard recruits under conditions of reward and fine. The direction of the differences in conformity is the same within reward and fine, with first born being *more* conforming than later born. The differences were significant, however, only in the reward conditions ($p = .06$). A comparison of first and later born persons on their conformity in the six experimental conditions of the study supports the overall finding. In six out of six groups, the first born are more conforming than the later born. This finding is directly opposite to our preceding findings from Experimental Situation A, but is quite consistent with the Ehrlich findings.

## DISCUSSION

We have presented a series of results which suggest that (*a*) first born females are more resistant to social influence than later born females and (*b*) first born males are less resistant to social influence than later born males. There is also a slight, but nonsignificant, indication that first born females have a higher need for achievement than later born females and that this relationship between birth order and need for achievement is stronger for the females than for the males.

Although the influence situations on which these data were obtained were different, we shall make the working assumption that there is sufficient comparability to permit us to draw certain conclusions about the important interactions between birth order, sex, and conformity.

In attempting to explain the original data relating birth order and conformity, Schachter (1959) made the assumption that the first born person who gets dethroned by later born sibs will be anxious about losing his position and losing the love of his parents, and would, thus, be oriented towards seeking attention, approval, and support from others. One could further assume that the first-and-only child—who by definition does not get dethroned by later sibs—is still involved in approval seeking behavior that has its source in the typically intense dependent relationship that parents form with an only child. The parents here may become the major source of support and approval, and anxiety over not pleasing them may become particularly intensified in the first-and-only child. Assuming this approval seeking, dependent behavior generalizes, one would expect the first born to conform more in a social influence situation than the later born. These assumptions are consistent with the Ehrlich data and with our data on males from Experimental Situation B, in which we find first born males less resistant (more conforming) to social influence than later born males.

However, we have reported another set of data—that involving the females from Experimental Situation A, which suggest that the first born female is more resistant to influence than the later born female. Thus, although the data from Experimental Situations A and B and our data and Ehrlich's are not strictly comparable, one could infer from the available data that the first born female is more independent and resistant to influence than the first born male. Accepting for the moment the validity of this inference, let us look at some possible assumptions one could make to explain it, as well as other research findings relevant to these assumptions.

One assumption relates early training in independence to the development of the need for achievement, and from this, to resistance to influence. Winterbottom (1958) reports a relationship between early training in independence and the development of need for achievement. Krebs (1958) offers further data on this point, although his correlation between independence training and n Ach was not significant. The picture becomes more complex when we look at findings relating the need for achievement with conformity. McClelland's (1953) analysis of the Asch data suggests a relation between high n Ach and independence from influence. Winterbottom reports that subjects high in n Ach less frequently

asked for help in a puzzle solving situation. Krebs finds that subjects with the most intense orientation towards achievement are the most resistant to opinion change. He also finds that there is greatest conformity when achievement is low and independence training is late. Similary, Walker and Heyns (1956 unpublished) find an inverse relation between achievement and conformity. This rather neat picture is somewhat complicated by the work of Samelson (1958) who finds that n Ach and social approach interacted with the experimental treatment, producing *higher conformity* under some conditions for persons high in the need for achievement. In support of this finding, Burdick (1955) suggests that the independent, achievement oriented person should conform in situations in which he perceives conformity as leading to achievement.

It seems fair to conclude that in general there is a positive relationship between early training in independence and the strength of need for achievement, but the apparent positive relationship between high n Ach and resistance to influence may be altered under certain conditions, as when conformity leads to achievement.

Applying these assumed relationships between early training in independence, development of the need for achievement, and resistance to influence to the interpretation of our data, we suggest that the first born female is involved in rearing later born sibs (Koch, 1955), and that this involvement gives the first born female more training in independence than the first born male. This independence training for the first born female could lead to a higher need for achievement and eventually to greater resistance in a social influence situation.

Another related assumption one could make involves the differential significance and timing of independence training for males and for females. The nature of the family situation is such that the young girl is being introduced to her adult role at an earlier age than the young boy. "Helping mother around the house" is more a part of the girl's later role than it is a part of the boy's later adult role. Here we are assuming that the first born male, although expected to assume some forms of responsibility in helping around the house, generally is not expected to be as responsible a person in these home activities as is the female, because these do not form an important part of his later adult role. His later responsibilities will lie within an occupational role in which he is still too young to be expected to form responsible behavior patterns. Some of these assumptions are supported by Koch's (1955) data in which she reports that generally girls are seen as more responsible than boys, and first born girls are rated higher in leadership (perhaps an indication of greater independence) than are first born boys. Koch, however, employed controls on sex of sibs and sib age separation, two conditions that were shown to interact significantly with birth order in effecting other variables such as leadership, ambition, etc.

We may now add to this picture another link in our chain of assumptions. Although the first born male and the first born female may be concerned over dependency and may be involved in seeking parental approval and love, the first born female initially has greater pressures than the male to develop responsibility as part of her role. She can handle both the needs to get parental approval and these pressures by the same activity: that is, by developing responsibility, she is receiving the approval of her parents. This does *not* mean that the male is not approved for exhibiting independent, responsible behavior. It does mean, however, that the parents expect greater responsibility at an earlier age from the female, apply greater pressures to the female to show it, and express greater approval when she exhibits such responsible, independent behavior. Consistent with this is Tuddenham's (1952) finding that in general girls are more approved by elders and peers than are boys.

What we are suggesting by our assumptions is that both the *significance* of the independence training and the *timing* of such training differ for the first born females as compared with first born males. Training in independence is more significant for the female and occurs at an earlier age than for the male.

Although one generally thinks of females as being more dependent than males, we are suggesting that the *first born* female is more independent and has a higher need for achievement than the *first born* male; and this independence results in greater resistance in a social influence situation on the part of the first born female as compared with the first born male. Without an analysis of birth order, the results from the work of Witkin, Lewis, Hertzman, Machover, Meissner, and Wapner (1954) which suggests that females are more field dependent than males, and the work of Crutchfield (1955) which suggests that females conform more than males, cannot be directly integrated into our own scheme. It is interesting to note, however that Crutchfield reports that with a sample of older females, he gets a whole range of conformity and resistance, with the general trend showing the

females to be *more* resistant to influence than the adult male sample. It is also interesting to note that in a replication of Witkin's work (Gruen, 1955), there is no significant difference between males and females on some of the tests on which Witkin et al. originally reported a difference. Thus, the picture that females are more dependent and conforming than males is itself more complex than originally thought, as some studies, including our own, suggest that some females are more resistant than some males. It would be of interest to reanalyze the data of Crutchfield, Witkin, and Gruen, controlling the apparently important variable of birth order.

The above set of assumptions provides a source for deriving additional hypotheses which are testable in a more systematic research project. For example, one would expect (*a*) the first born female to have greater responsibility demands from her parents than the first born male, (*b*) parents to perceive early responsibility as being more important for the first born female child than for the first born male child, and (*c*) the first born female to be more resistant to social influence than the first born male. We have inferred support for this third hypothesis from our data. However, a strict test of that hypothesis would involve placing males and females into the *same* influence situation and measuring their resistance to influence.

## SUMMARY

The combined results from three separate studies using three separate samples of subjects suggest the following three conclusions: first born persons have a higher need for achievement than later born persons; first born females exhibit greater resistance to influence than later born females; and first born males exhibit less resistance to influence than later born males.

These findings are taken to be consistent with a set of assumptions that the first born female is more significantly involved in independence training than the first born male. This early independence training produces a greater need for achievement and leads to greater resistance to influence for the first born females. First ordinal position for the male, on the other hand, produces greater affiliative dependency and leads to greater conformity in an influence situation.

## REFERENCES

ADLER, A. *Social interest: A challenge to mankind.* (Trans. by J. Linton & R. Vaughan) London: Faber & Faber, 1945.

ATKINSON, J. W. (Ed.) *Motives in fantasy, action, and society.* Princeton: Van Nostrand, 1958.

ATKINSON, J. W., & LITWIN, G. H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *J. abnorm. soc. Psychol.*, 1960, **60**, 52–63.

BURDICK, H. A. The relationship of attraction, need achievement, and certainty to conformity under conditions of simulated group atmosphere. Unpublished doctoral dissertation, University of Michigan, 1955.

CRUTCHFIELD, R. S. Conformity and character. *Amer. Psychologist*, 1955, **10**, 191–198.

FRENCH, E. G. Development of a measure of complex motivation. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society.* Princeton: Van Nostrand, 1958. Pp. 242–248.

GRUEN, A. The relation of dancing experience and personality to perception. *Psychol. Monogr.*, 1955, **69**(14, Whole No. 399).

KOCH, H. L. Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children. *Genet. psychol. Monogr.*, 1955, **52**, 3–50.

KREBS, A. M. Two determinants of conformity: Age of independence training and achievement. *J. abnorm. soc. Psychol.*, 1958, **56**, 130–131.

MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive.* New York: Appleton-Century-Crofts, 1953.

MANDLER, G., & COWEN, J. E. Test anxiety questionnaires. *J. consult. Psychol.*, 1958, **22**, 228–229.

SAMELSON, F. The relation of achievement and affiliation motives to conforming behavior in two conditions of conflict with a majority. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society.* Princeton: Van Nostrand, 1958. Pp. 421–433.

SCHACHTER, S. *The psychology of affiliation.* Stanford, Calif.: Stanford Univer. Press, 1959.

TUDDENHAM, R. D. Studies in reputation: I. Sex and grade differences in school children's evaluation of their peers. *Psychol. Monogr.*, 1952, **65**(1, Whole No. 333).

WINTERBOTTOM, M. R. The relation of need for achievement to learning experiences in independence and mastery. In J .W. Atkinson (Ed.), *Motives in fantasy, action, and society.* Princeton: Van Nostrand, 1958. Pp. 453–478.

WITKIN, H. A., LEWIS, HELEN B., HERTZMAN, M., MACHOVER, KAREN, MEISSNER, PEARL B., & WAPNER, S. *Personality through perception.* New York: Harper, 1954.

(Received January 23, 1961)

*Journal of Abnormal and Social Psychology*
1962, Vol. 64, No. 2, 160–162

# THE RECALL OF DREAMS:

## ITS RELATION TO REPRESSION AND COGNITIVE CONTROL

FRANK M. LACHMANN, BENJAMIN LAPKIN, AND NORMAN S. HANDELMAN [1]

*Veterans Administration Regional Office, New York*

With the publication *The Interpretation of Dreams*, Freud (1959) opened the way for a more scientific approach to the study of dream phenomena. The psychoanalytic investigations which followed, and the inferences drawn, were based primarily upon individual case histories, difficult to replicate. Experimental studies have been for the most part nontheoretical in nature; Ramsey (1953) summarizes this voluminous empirical literature. Until recently the psychoanalytic and the experimental approaches were divergent paths, overlapping only accidentally. The experimental findings of Kleitman (1960), demonstrating the universality of the dream process, have provided a baseline for laboratory investigation along theoretical psychoanalytic lines. Dement and Kleitman (1957) and Goodenough, Shapiro, Holden, and Steinschriber (1959) noted the occurrence of rapid conjugate eye movements and electroencephalographic activity at specific intervals during sleep which were found to coincide with dream activity. Even subjects who said they had never remembered dreaming prior to the study reported dreams when awakened during eye movement periods.

Dreaming seems clearly a universal phenomenon, an "intrinsic part of normal sleep" (Dement & Kleitman, 1957, p. 345). It further appears that there is a great uniformity in the pattern of dream activity among people. One would expect a comparable uniformity in the ability to remember dreams and yet, as the literature clearly shows, this is not the case. There are individuals who scarcely recall dreaming while others report dream activity almost every night. This has led to investigations to account for the differences between such divergent types of behavior. Schonbar (1959) found that subjects who recalled dreams more often than 25% of the time (the high dream recall group) were more "intelligent" and experienced more "overt anxiety" than the nonrecall group (those who reported dreams once a month or less). In addition, she reported a high correlation between remembering dreams and the contentless recall of dreaming. From this she hypothesized that low dream recall is related to repression, which conforms to the general psychoanalytic theory (Freud, 1959) that forgetting dreams or failing to recall dreams is a function of repression.

Personality differences associated with the variations in the amount of dream recall may thus be a reflection of differences among individuals in the utilization of repression. Other correlates of repression have been derived from the research in cognitive controls (Gardner, Holzman, Klein, Linton, & Spence, 1959) which has demonstrated that the leveling-sharpening tendency is related to the operation of the ego defenses, specifically repression. Holzman and Gardner (1959) compared scores on the Schematizing test (described below) with ratings of repression, based on Rorschach test indicators, e.g., few responses, vagueness, emphasis on color. All extreme repressers were also extreme levelers although the converse was not found.

The studies of Schonbar (1959), and Holzman and Gardner (1959) suggest that dream recall, repression and the leveling-sharpening continuum are associated. It is the hypothesis of the present study that subjects who recall few if any dreams, i.e., who use repression as a major defense, will show the leveling tendency. It is further hypothesized that subjects who recall many dreams, i.e., do not rely upon repression as their major defense,will show the sharpening tendency.

The intention here is to relate empirically two conceptual frames of reference, the psychoanalytic concept of repression and the cognitive controls concept of leveling-sharpening, with respect to the recall of dreams.

### METHOD

The subjects were students in two introductory psychology classes at a municipal college. The tests utilized were administered by one of the authors. Since he was also the instructor of the

[1] The authors are indebted to Lloyd Silverman for his suggestions in formulating this study and to Edward Bryant for his assistance with the apparatus.

Frank M. Lachmann and Norman S. Handelman are now associated with the Postgraduate Center for Psychotherapy, New York. Benjamin Lapkin is now at Hillside Hospital, New York.

classes, it was felt that at the time of testing sufficient rapport had been established for cooperation. Material distributed was numbered to permit subsequent collation. The subjects were assured of anonymity and that nonparticipation would in no way affect their class standing. The two classes totaled 52 students, 11 males and 41 females. The mean age of the group was 19.7 years with a range from 17.5 to 42.7 years. Six of the group were married. No student had received psychotherapy or psychoanalysis.

### *Recording of Dreams*

Following Schonbar (1959), booklets were prepared in which the subjects were to record daily their dream activity for 14 consecutive days. Space was provided to record the content of the dreams or to note that no dream was recalled. The distribution of booklets to the subjects was accompanied by instructions identical to those published by Schonbar. While she collected the dream reports daily for 4 weeks, the present procedure involved a 2-week period, at the end of which the entire booklet was collected. On alternate days, the subjects were reminded to continue recording in the booklets. The measure of dream recall was the number of days in which subjects recorded at least one dream. The sample was divided into quartiles based upon this measure, the upper quartile being the high recall group and the lowest quartile, the low recall group.

### *Schematizing Test*

One week after the distribution of the dream report booklets, the Schematizing test was administered to each class as a group. Subjects were told that they were taking a test of perceptual accuracy and no association was made with the dream study.

A detailed description of the apparatus and administration of this test appears in Gardner et al. (1959). The test required subjects to estimate sizes of squares projected individually onto a screen. There were 10 series of 15 squares each, the first 5 squares of each series in an ascending order and the subsequent 10 squares of the same sizes repeated in a random order. Each following series began by dropping the smallest square of the preceding series and ascending one step. The smallest square in the first series was 1.2 inches and the largest square in the tenth series was 13.7 inches. The 150 slides were presented automatically at a uniform exposure of 3 seconds, with an 8-second interval between exposures. An electromagnetic shutter, operated by Hunter timers, controlled the presentations of an automatic load 35-millimeter slide projector.

As described by Holzman and Gardner (1959), the Schematizing test was scored for both ranking accuracy and degree of lag. A composite score was obtained for each subject based upon this double criterion. Subjects were classified as levelers if they scored in the upper third and sharpeners if they scored in the lower third of the distribution.

### *Subjects*

Both parts of the study were completed by 38 subjects, 8 males and 30 females. The remainder either failed to return the dream report booklets, returned the booklets with an indication that they did not participate, or were absent when the Schematizing test was administered. In view of sex differences in the Schematizing test noted by Gardner et al. (1959), and in consideration of the small number of men in the sample, only females were included in the following analysis of data. The final group evaluated consisted of 30 women, ranging in age from 17.7 to 42.7 years, with a mean age of 19.9. Five were married.

## RESULTS

Tabulation of the dream recall reports indicated a range of days on which dreams were recorded from 0 to 10 with a median of 4 days. The upper quartile consisted of 9 subjects who reported dreams on 6 days or more, the lowest quartile of 7 subjects who reported dreams on no more than 1 day and the middle group consisted of 14 subjects who reported dreams on 2–5 days. A classification of subjects on their Schematizing test scores yielded 11 levelers, 10 sharpeners, and a midgroup of 9 subjects. Table 1 presents the number of levelers and sharpeners in each of the dream recall categories.

An analysis of variance was computed, comparing the subjects in the high, medium, and low dream recall groups with respect to their Schematizing test scores. Table 2 presents this analysis.

An $F$ value of 3.31 was obtained which approximates the $F$ value of 3.36 required for significance at the .05 level. In view of this finding a $t$ test was computed between the two groups which showed the greatest difference between their Schematizing test scores: the medium and high dream recall groups ($t = 2.67$, $p$ between .01 and .02). A $t$ test was also computed be-

TABLE 1

NUMBER OF LEVELERS AND SHARPENERS IN THE DREAM RECALL GROUPS

| Schematizing test category | Dream recall groups | | |
|---|---|---|---|
| | High | Medium | Low |
| Levelers | 1 | 7 | 3 |
| Sharpeners | 5 | 2 | 3 |
| Unclassified | 3 | 5 | 1 |

TABLE 2

ANALYSIS OF VARIANCE OF SCHEMATIZING TEST SCORES

| Source of variation | *SS* | *df* | *MS* | *F* |
|---|---|---|---|---|
| Between groups | 1079.33 | 2 | 539.67 | 3.31 |
| Within groups | 4406.64 | 27 | 163.21 | |
| Total | 5485.97 | 29 | | |

tween the medium and low dream recall and a *t* value of 1.15, below an acceptable level of significance, was obtained. This indicates that high dream recallers tend to score toward the sharpening end of the continuum. Subjects in the medium dream recall group tend to score toward the leveling end. No definite results were obtained with regard to the low dream recall group. Inspection of the data of this latter group cited in Table 1 reveals the presence of three levelers and three sharpeners.

To evaluate the overall relationship between number of dreams and the leveling-sharpening continuum a Pearson product-moment correlation coefficient was computed. A nonsignificant correlation of .20 was obtained. Thus though high dream recall is associated with the sharpening tendency and medium dream recall with leveling, it cannot be substantiated at this point that low dream recall fits into this continuum.

## DISCUSSION

The findings of this study point to a difference between the medium and high dream recall groups with respect to cognitive controls. High dream recall is associated with the cognitive control of sharpening, thus supporting one of the hypotheses of this study. It had been hypothesized initially that the converse would hold also, namely, that the low dream recallers would show the leveling tendency. However, the results do not support this contention, indicating instead that low dream recallers actually are at either extreme of the schematizing continuum. It thus appears that the repression-leveling relationship is not supported. This is contrary to the findings of Holzman and Gardner (1959) and Gardner et al. (1959) who reported a significant association along these lines. To understand these divergent findings the following possibilities are presented. First, different criterion measures were used: Holzman and Gardner used Rorschach protocols to measure reliance upon repression while the present investigators used the amount of dream recall. Furthermore, the absence of normative criteria for leveling and sharpening on the Schematizing test makes interstudy comparisions difficult. Further investigation is needed to provide normative data in this area.

## SUMMARY AND CONCLUSIONS

The present study investigated the relationship between the recall of dreams and the cognitive controls continuum of leveling-sharpening. Thirty college students reported dream activity daily for a 14-day period and the amount of dream recall was related to their scores on the Schematizing test. The results confirmed the hypothesis that subjects who recall many dreams, i.e., who do not rely primarily upon repression, would show the sharpening tendency. A second hypothesis, postulating a relationship between low dream recall and the leveling tendency, was not supported.

It appears that the psychoanalytic concept of repression can be associated with the theory of cognitive controls, in that there is a statistical relationship between repression as measured by the extent of dream recall and the leveling-sharpening tendency as measured by the Schematizing test.

## REFERENCES

DEMENT, W., & KLEITMAN, N. The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming. *J. exp. Psychol.*, 1957, **53**, 339–346.

FREUD, S. *The interpretation of dreams.* (Originally published 1900) New York: Basic Books, 1959.

GARDNER, R., HOLZMAN, P. S., KLEIN, G. S., LINTON, HARRIET, & SPENCE, D. P. Cognitive control: A study of individual consistencies in cognitive behavior. *Psychol. Issues*, 1959, **1**, No. 4.

GOODENOUGH, D. R., SHAPIRO, A., HOLDEN, M., & STEINSCHRIBER, L. A comparison of "dreamers" and "nondreamers": Eye movements, electroencephalograms, and the recall of dreams. *J. abnorm. soc. Psychol.*, 1959, **59**, 295–302.

HOLZMAN, P., & GARDNER, R. Leveling and repression. *J. abnorm. soc. Psychol.*, 1959, **59**, 151–155.

KLEITMAN, N. Patterns of dreaming. *Scient. American*, 1960, **203**, 82–88.

RAMSEY, G. V. Studies of dreaming. *Psychol. Bull.*, 1953, **50**, 432–455.

SCHONBAR, ROSALEA A. Some manifest characteristics of recallers and nonrecallers of dreams. *J. consult. Psychol.*, 1959, **23**, 414–418.

(Received January 30, 1961)

# *Psychology and Mental Health*

A report of the Institute on Education and Training for Psychological Contributions to Mental Health, held at Stanford University in August, 1955.

*Edited by* CHARLES R. STROTHER

Topics discussed at the Institute:

Training Needs of Psychologists in Community Mental Health

Specialization in Training

Practicum Training

Training for Therapy

Training for Research in the Mental Health Field

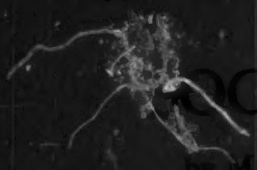
Problems of Departmental Organization

Price: $1.75

*Order from:*

American Psychological Association, Inc.
1333 Sixteenth Street, N. W.
Washington 6, D. C.

*Yale Studies in Attitude and Communication, 4*



# SOCIAL JUDGMENT

by Muzafer Sherif and Carl I. Hovland

This process of attitude modifications has been studied by the members of the Yale Communication and Attitude Change Program through systematic investigation of a number of variables. The present volume presents a new approach to the problems of attitude and opinion change. Convinced that attitude research should be solidly based on previous work in psychological processes, particularly of judgment, the authors begin by examining earlier research on the judgment of simple stimulus material under well-controlled laboratory conditions. The more complex problems of judgment of social issues are investigated, using new studies involving judgment of the social position of Negroes, the prohibition issue, and the choice of presidential candidates. $6.00

*Volume 1*
THE ORDER OF PRESENTATION IN PERSUASION
*by Hovland, Mandell, Campbell, Brock, Luchins, Cohen, McGuire, Janis, Feierabend, and Anderson.* $5.00

*Volume 2*
PERSONALITY AND PERSUASIBILITY
*by Janis, Hovland, Field, Linton, Graham, Cohen, Rife, Abelson, Lesser, and King.* $6.00

*Volume 3*
ATTITUDE ORGANIZATION AND CHANGE
*by Rosenberg, Hovland, McGuire, Abelson, and Brehm.* $5.00

*Yale Pepperounds*
CHILD TRAINING AND PERSONALITY *by John W. M. Whiting and Irvin L. Child* $1.75

RECEPTORS AND SENSORY PERCEPTION *by Ragnar Granit* $1.95

FRUSTRATION AND AGGRESSION *by Dollard, Doob, Miller, Mowrer, and Sears* $1.25

THE INTEGRATIVE ACTION OF THE NERVOUS SYSTEM *by Sir Charles Sherrington* $1.95

*Yale University Press* Y A L E *New Haven and London*
*Canadian orders: McGill University Press, Montreal 2*